Journal
of
The Franklin Institute
Devoted to Science and the Mechanic Arts

Vol. 270 JULY, 1960 No. 1

INTRODUCTION *

BY

JOSEPH GRAY JACKSON [1]

There is a considerable analogy between doing research and building a house. For centuries, most houses were built entirely at the site, without employing the services of an architect.

Much research and development has been done like this—and good work, too. Donald Douglas reports that when he first entered the aircraft industry, no drawings were being made and the parts for constructing a plane were whittled out in the shop. R. G. LeTourneau developed his earth-moving equipment with a welding torch in hand.

Nowadays, however, the services of an architect are usually employed in building a house, and similarly in research, the tendency is to emphasize the planning stage. Many people, however, still launch into a research program either without any literature search, or with a very meager effort in this direction.

Anyone concerned with research direction should ask himself the question, "Can my company afford the luxury of discovering things again which others have discovered two or three times before?"

The difficulty in making an extensive literature search in many fields is so great as to provide an excuse for the inadequacy of the effort in this direction. Why not employ the computer to do this work, with the advantages of instantaneous answers, tremendous scope and freedom from human error?

* The four papers comprising this symposium were presented at the Tuesday morning session of the Conference on Scientific Communication, as part of the Annual Meeting of the American Association for the Advancement of Science, held December 29 and 30, 1959, in Chicago, Illinois.

[1] William Steell Jackson and Sons, Philadelphia, Pa. Mr. Jackson presided over the session on "Research in Documentation."

(Note—The Franklin Institute is not responsible for the statements and opinions advanced by contributors in the JOURNAL.)

The field of organic chemistry is a "natural," because precise architect's drawings can be made, known to the chemist as structural formulae, which will infallibly call for a particular compound if it exists in the literature. Two of the papers in this symposium are based upon the experience in the United States Patent Office in mechanical searching of organic chemistry.

A somewhat different problem is presented in searching literature where the subject matter sought cannot be embodied in an architect's drawing, but instead involves the presence of selected features or characteristics not structurally interrelated. A valiant effort to apply the computer to searching in this area has been undertaken by the American Society for Metals in cooperation with Western Reserve University in the metallurgical field, and two of the symposium papers are drawn from this experience.

In the future, an extensive field of activity for scientific societies, universities and research foundations will be the encoding, computer storage, and retrieval of scientific information in various fields.

SEMI-AUTOMATIC INDEXING AND ENCODING *

BY

JULIUS FROME [1]

INTRODUCTION

Most mechanized information-retrieval systems are dependent upon the extraction of information from documents, processing such information in some manner and finally placing this information in a file so that the information can be retrieved by machine. Any such retrieval system is only as good as the degree of accuracy and thoroughness of the extraction of information from the document.

Analysis can be defined as the process of extracting information from a document. This information is often in the form of notions or concepts.

Indexing processes the information resulting from analysis by placing it in lists.

Coding creates and assigns specific unambiguous terms as substitutes for such notions or concepts (13).[2] The terms may come from the document itself.

Encoding prepares a document for storage by substitution of codes or terms for the information already extracted from the document.

To date, in the development of most of the mechanized information-retrieval systems, including the three mechanical retrieval systems developed in the U. S. Patent Office (1, 2, 3) the major portion of the effort of analyzing and encoding documents has been essentially manual. Only a small portion of this work has been mechanized.

It is desirable to attempt to shift the major burden of the analyzing and encoding from humans to machines. This paper will describe and illustrate an experimental attempt to do so. This attempt suggests a possible interim solution to these problems. The extraction, listing, indexing and encoding of the structural formulae not specifically named in the document are treated in another manner not included in this paper.

Four extremely important factors which have been considered are those of (1) personnel, (2) training, (3) product obtained and (4) cost. I have called this "a semi-automatic indexing and encoding method." In this experiment humans are used for what they are best suited,

* Adapted from Patent Office Research and Development Reports No. 17.

[1] Patent Research Specialist, U. S. Patent Office, Washington, D. C.

[2] The boldface numbers in parentheses refer to the references appended to this paper.

for example, reading and recognition of concepts, and machines for what they can do best, for example, fast, accurate, and repetitive action.

The usual two-step procedure was used, consisting, first, of extracting the material from the document and, then, encoding it in a form suitable for making searches. The particular method used in the first step additionally resulted in a method for indexing. Such auxiliary indexing is of extreme importance in serving as a surrogate for the document itself for the purpose of browsing in manual retrieval systems. The first step will, accordingly, be called indexing.

An excellent start in automatic abstracting has been made by H. P. Luhn (**4, 5, 6, 7**) and associates. However, these systems cannot now be used by the Patent Office. Even if the text of the documents were available in "machine-readable" form and a computer such as the IBM 704 were available, no program for performing the type of analysis we find necessary is now known. The ultimate solution awaits not only the development of machines able to read any text in any type font, but the discovery of those rules of language that would enable machines to process the text on the basis of its information content. Until such developments are forthcoming, humans will still play an important part in the task of analyzing and encoding documents. H. P. Luhn (**8**) and C. K. Schultz *et al.* (**10**) also suggests mechanized methods for indexing terms which they extract.

V.2
196

INDEXING

Indexing is the listing of words, subjects, or concepts found in, or otherwise associated with a document, which enables a person to find that document. Indexing in depth is an extremely useful tool for Patent Office use. Deep indexing by the usual manual methods is quite expensive and has required very highly trained personnel (**11**). In "subject" indexing particularly, extracted terms or concepts are placed under headings according to some specified scheme of classification. This requires personnel skilled both in the discipline involved and in the art of classification.

Most indexes carry no indication of the depth of the indexing, since this factor is left to the judgment of the indexer. While he is always conscious of the time required to make the index, he is, perhaps, not as conscious of all the varied requirements of the users. If a method of indexing is chosen which has definite rules, a less skilled indexer can be used. The uniform application of these rules would control the depth of indexing, and knowledge of them will indicate that depth. Our experimental method enables the user to know both the depth of indexing and the point of view from which the indexing was carried out.

Indexes in which the user can find the item desired either in the *language of the document* or under an appropriate subject heading appear desirable.

The following steps were taken in preparing the index:

1. Determination of the point of view or objective of indexing, for example, extracting the chemical compounds from a document or chemical processes; or, in disciplines other than chemistry, extracting electrical components or systems.
2. Determination of depth of indexing desired, for example, extracting *all* chemical compounds, *all* steroid compounds, *only* phosphate compounds or *only* electrical components.
3. Determination of the rules for extracting terms or concepts, in accordance with the point of view and depth of indexing, is communicated to the indexer. The more precise and simple the rules are, the less training of personnel is needed, and the better the index.
4. Conversion of extracted terms to machinable form, for example, converted into punched cards.
5. Elimination and counting of duplicates by machine processing. The elimination of synonyms is not done at this time.
6. Printing of lists or indexes by machine.

These first steps now result in a list of terms for each document. The machine can then arrange and print an alphabetical list of all the terms in all the documents indexed, again eliminating the duplicates, while indicating the plural documents in which each term appears.

Experimental work has been carried on in the U. S. Patent Office since April, 1959. Several specific examples of this system will now be described.

Steroid Example

All "steroid compounds," not shown by structural formulae, are extracted from each patent of the 2,100 patents in the twelve steroid subclasses (1).

The technical personnel were instructed to underline each word which recited a steroid compound as it occurred and *every time* it occurred. Thus a steroid compound occurring seven times was underlined seven times. Hence no effort or time was lost in determining whether a compound was previously underlined. Structural formulae were not underlined. The underlined patent was next given to a punch card operator who punched the underlined work in alphanumeric text. One card was punched for each steroid compound. Trailer cards were used as needed. The patent number in which the word occurred was also punched in each card. A RAMAC 305, by suitable program, accepted these cards, copied each card into its disk memory and then eliminated duplicate information. Next it punched out a new deck of cards which eliminated duplicates and printed a list of steroid compounds in alphabetical order, containing no duplicates. Each card and its listing included the patent number and the number of times

the term appeared in the patent. This list (see Table I) constitutes an index of *all* the steroid compounds in each patent, excluding those represented by structural formulae. These could be easily included if desired. The resulting list tends to be complete and accurate because of the way it is generated. In most patents each steroid compound is usually repeated two or more times. If the indexer failed to underline one occurrence, it is probable that the second or third occurrence would

TABLE I.

A partial listing of all the steroid compounds in Patent No. 2,905,678. The two-digit number is one less than the number of times the compound occurred in the patent.

2905678	6 ACETATE 4 PREGNENE	00
2905678	5A 6B DIHYDROXY 3 KETO PREGNANE	00
2905678	5A 6B DIHYDROXY 3 KETO PREGNANES	00
2905678	5A 6B DIHYDROXY PREGNANE	00
2905678	5A 6B DIHYDROXY PREGNANES	01
2905678	5A 6B DIHYDROXY 3 11 20 TRIKETO PREGNANE	00
2905678	5A 6B DIHYDROXY 3 11 20 TRIKETO PREGNANE 6 ACETATE	00
2905678	5A 6B DIHYDROXY 3 11 20 TRIKETO PREGNANE 6 FORMATE	00
2905678	5A HYDROXY 6 ACETOXY PREGNANE	00
2905678	5A HYDROXY 6 ACETOXY PREGNANES	00
2905678	5A HYDROXY 6 FORMOXY PREGNANE	00
2905678	5A HYDROXY 6B ACETOXY PREGNANE	00
2905678	5A HYDROXY 6B FORMOXY PREGNANE	00
2905678	5A HYDROXY 6B FORMOXY PREGNANES	01
2905678	5A 6B 17A TETRAHYDROXY 3 11 20 TRIKETO PREGNANE 6 FORMATE 21 ACETATE	00
2905678	5A 6B 17A 21 TETRAHYDROXY 3 11 20 TRIKETO PREGNANE 21 ACETATE	02
2905678	5A 6B 17A 21 TETRAHYDROXY 3 11 20 TRIKETOPREGNANE 21 ACETATE	01
2905678	5A 6B 17A TRIHYDROXY 3 11 20 TRIKETO PREGNANE	00
2905678	5A 6B 17A TRIHYDROXY 3 11 20 TRIKETO PREGNANE 6 FORMATE	03
2905678	5A 6B 21 TRIHYDROXY 3 11 20 TRIKETO PREGNANE 6 FORMATE 21 ACETATE	03
2905678	5A 6B 21 TRIHYDROXY 3 11 20 TRIKETO PREGNANE 6 FORMATE 21 BENZOATE	02
2905678	5A 6B 21 TRIHYDROXY 3 11 20 TRIKETO PREGNANE 6 FORMATE 21 BUTYRATE	00
2905678	5A 6B 21 TRIHYDROXY 3 11 20 TRIKETO PREGNANE 6 FORMATE 21 CAPROATE	00
2905678	5A 6B 21 TRIHYDROXY 3 11 20 TRIKETO PREGNANE 6 FORMATE 21 HEMISUCCINAT	00
2905678	,E	00
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 21 DIACETATE	01
2905678	6B 17A DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE	05
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE 21 ACETATE	09
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE 21 BENZOATE	03
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE 21 BUTYRATE	01
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE 21 CAPROATE	00
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE 21 ESTERS	00
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE 21 HEMISUCCINATE	01
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE 21 PHENYLACETATE	00
2905678	6B 32	00
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE 21 ISOVALERATE	01
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE 21 PHENYLACETATE	00
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE 21 PROPIONATE	01
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE 21 T BUTYLACETATE	01
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE 21 TRICARBALLYLAT	00
2905678	,E	00
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 6 FORMATE 21 VALERATE	01
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 21 A	00
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 21 ACETATE	06
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 21 BENZOATE	04
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 21 T BUTYLACETATE	00
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 21 T BUTYL ACETATE	01
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 21 BUTYRATE	02
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 21 HEMISUCCINATE	02
2905678	6B 21 DIHYDROXY 3 11 20 TRIKETO 4 PREGNENE 21 PHENYLACETATE	00

United States Patent Office 2,870,190 Patented Jan. 20, 1959

1

2,870,190

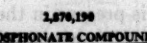
PHOSPHONATE COMPOUNDS

Bill F. Burgett and Henry Tolbmith, Midland, Mich., assignors to The Dow Chemical Company, Midland, Mich., a corporation of Delaware

No Drawing. Application July 3, 1957 Serial No. 669,718

5 Claims. (Cl. 260—461)

The present invention relates to phosphonate compounds having the formula



In this and succeeding formulas, X_1 represents X_2 or hydrogen and X_2 represents a radical having the formula

wherein R is a lower alkyl radical. The expression "lower alkyl" is employed in the present specification and claims to refer to the alkyl radicals containing from 1 to 4 carbon atoms, inclusive. These new compounds are very viscous liquids which are readily soluble in water and somewhat soluble in many organic solvents. They have been found to be useful as parasiticides and are adapted to be employed as active toxic constituents of dust and liquid compositions for the control of many insect and fungal pests such as *Alternaria solani, Fusarium solani, Pythium spp.* and *Rhizoctonia solani.*

The new compounds may be prepared by causing a reaction between ethylenediamine, formaldehyde and a dialkyl phosphite of the formula

Where it is desired to introduce two $-CH_2PO(OH)(OR)$ groups into the molecule, one molecular proportion of ethylenediamine is employed with two molecular proportions of formaldehyde and two molecular proportions of dialkylphosphite. Where it is desired to introduce four $-CH_2PO(OH)(OR)$ groups into the molecule, one molecular proportion of ethylenediamine is employed with four molecular proportions of each of the formaldehyde and dialkyl phosphite reactants. Although the details of the reaction mechanism are not completely understood, the reaction takes place smoothly at the temperature range of from 0° to 50° C. with the formation of the desired product and alkanol of reaction. The reaction is somewhat exothermic and the temperature may be controlled by regulating the rate of contacting the reactants and by external cooling. Upon completion of the reaction, the reaction mixture is fractionally distilled under reduced pressure to separate low boiling constituents and obtain the desired product as a liquid residue.

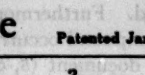
In carrying out the reaction, a mixture of ethylenediamine and dialkyl phosphite is contacted portionwise with an aqueous solution of formaldehyde. This operation is carried out with stirring and at a temperature of from 0° to 50° C. Upon completion of the reaction, the reaction mixture is fractionally distilled under reduced pressure at gradually increasing temperatures up to a temperature of 150° C. to separate low boiling constituents

2

and obtain the desired product as a liquid residue. Since the desired products are somewhat unstable at temperatures in excess of 150° C., exposure to such elevated temperatures for any appreciable period should be avoided.

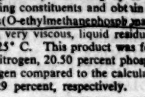
The following examples illustrate the invention but are not to be construed as limiting:

Example 1.—Ethylenediamine N,N'-bis(O-ethyl methanephosphonate).



A 37 percent aqueous solution of formaldehyde containing 16.5 grams (0.55 mole) of formaldehyde was added portionwise with stirring to a mixture of 15 grams (0.25 mole) of ethylenediamine and 75.9 grams (0.55 mole) of O,O-diethylphosphite. The addition was carried out in 15 minutes and at a temperature of from 25° to 35° C. Following the addition, stirring was continued for 40 minutes, and the reaction mixture thereafter diluted with 70 grams of methylcyclohexane. The resulting mixture was fractionally distilled under reduced pressure at temperatures gradually increasing up to 150° C. to remove low boiling constituents and obtain as ethylenediamine N,N'-bis(O-ethylmethanephosphonate) product as a dark brown, very viscous, liquid residue having a density of 1.4 at 25° C. This product was found to contain 9.27 percent nitrogen, 20.50 percent phosphorus and 7.09 percent hydrogen compared to the calculated values 9.21, 20.40 and 7.29 percent, respectively.

Example 2.—Ethylenediamine N,N,N',N'-tetra (O-ethyl methanephosphonate).



A 37 percent aqueous solution of formaldehyde containing 24 grams (0.80 mole) of formaldehyde was added portionwise with stirring to a mixture of 9 grams (0.15 mole) of ethylenediamine and 109.1 grams (0.80 mole) of O,O-diethyl phosphite. The addition was carried out in 15 minutes and at a temperature of from 20° to 30° C. Following the addition, stirring was continued for 90 minutes to complete the reaction. The reaction mixture was then processed as described in Example 1 to obtain an ethylenediamine N,N,N',N'-tetra(O-ethyl methanephosphonate) product as a light tan, viscous, liquid residue having a density of 1.2 at 25° C. This product contained 5.49 percent nitrogen and 20.86 percent phosphorus compared to the theoretical values of 5.11 percent and 22.16 percent, respectively.

Example 3.—Ethylenediamine N,N'-bis(O-isopropyl methanephosphonate).

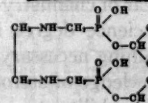


FIG. 1. An example of the underlining in a phosphate patent.

be underlined. Furthermore, since the list also contains the number of times the compound occurred, this may be an indication of its importance in the document (5, 6). This list can be attached to the patent to aid in manual searching of the patent. It also enables one to see quickly whether a particular steroid compound is present in the patent. Specific examples of an underlined patent and such a prepared list can be seen in Fig. 1 and Table II. A flow chart of the program is shown in Figs. 2 through 6. Another RAMAC program takes the alphabetized

TABLE II.—*Listing in a Phosphate Patent.*

2870190	DIALKYL PHOSPHITE	01
2870190	DIALKYLPHOSPHITE	00
2870190	DIALKYL PHOSPHITE REACTANTS	00
2870190	ETHYLENEDIAMINE N N& BIS-O ETHYL METHANEPHOSPHONATE-	00
2870190	ETHYLENEDIAMINE N N& BIS-O ETHYLMETHANEPHOSPHONATE	00
2870190	ETHYLENEDIAMINE N N N& N& TETRA-O ETHYLMETHANEPHOSPHONATE	00
2870190	ETHYLENEDIAMINE N N N& N& TETRA-O ETHYLMETHANEPHOSPHONATE- PRODUCT	00
2870190	ETHYLENEDIAMINI N N& BIS-O ISOPROPYL METHANEPHOSPHONATE	00
2870190	ETHYLENEDIAMINE N N N& N& TETRA-O NORMALBUTYL METHANEPHOSPHONATE	02
2870190	ETHYLENEDIAMINE N METHYLPHOSPHONATES	00
2870190	ETHYLENEDIAMINE N N& BIS-O METHYL METHANEPHOSPHONATE	00
2870190	ETHYLENEDIAMINE N N& BIS-O ISOBUTYL METHANEPHOSPHONATE	00
2870190	ETHYLENEDIAMINE N N N& N& TETRA-O METHYL METHANEPHOSPHONATE-	00
2870190	ETHYLENEDIAMINI N N N& N& TETRA-O NORMALPROPYL METHANEPHOSPHONATE	00
2870190	ETHYLENEDIAMINE N N N& N& TETRA-O ETHYL METHANEPHOSPHONATE	01
2870190	ETHYLENEDIAMINE N N& BIS-O ETHYL METHANEPHOSPHONATE	00
2870190	ETHYLENEDIAMINE N N& BIS-O ISOPROPYL METHANEPHOSPHONATE	00
2870190	O O DIETHYL PHOSPHITE	00
2870190	O O DINORMALBUTYL PHOSPHITE	00
2870190	O O DIMETHYL PHOSPHITE	00
2870190	O O DIISOBUTYL PHOSPHITE	00
2870190	O O DINORMALPROPYL PHOSPHITE	00
2870190	O O DI-LOWER ALKYL-PHOSPHITE	00
2870190	O O DIISOPROPYL PHOSPHITE	00
2870190	PHOSPHONATE COMPOUNDS	00
2870190	PHOSPHONATE COMPOUND	00

deck of the entire file of documents, and assigns address numbers of the RAMAC storage (9) for each different concept.

Personnel

In contrast to other methods, this method of indexing was done with personnel who have only a B.S. degree in chemistry. No other formal schooling appeared to be necessary. In fact, in two instances, personnel with only 1 year of chemistry were trained to do this work satisfactorily. However, a science degree is more desirable because it reduces the amount of training necessary. Highly trained indexers are not needed because a knowledge of steroid chemistry is not necessary. All that is required is the ability to recognize that a word describes a steroid compound.

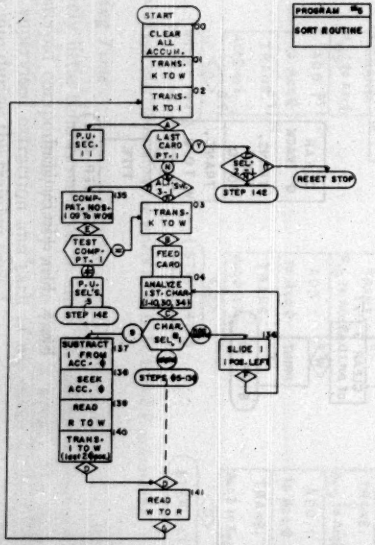


FIG. 2. Flow chart used on the RAMAC 305 for eliminating the duplicates.

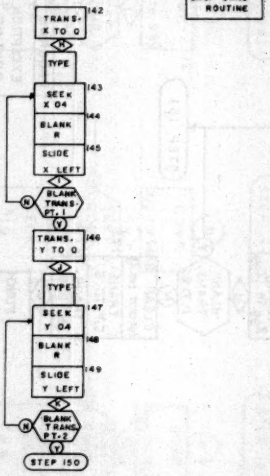


FIG. 3. Flow chart continued.

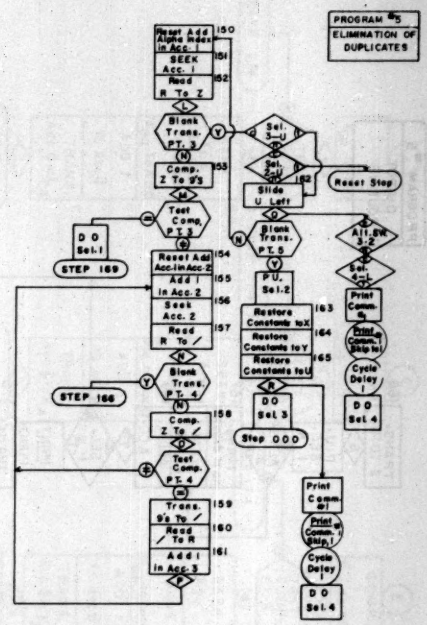


FIG. 4. Flow chart continued.

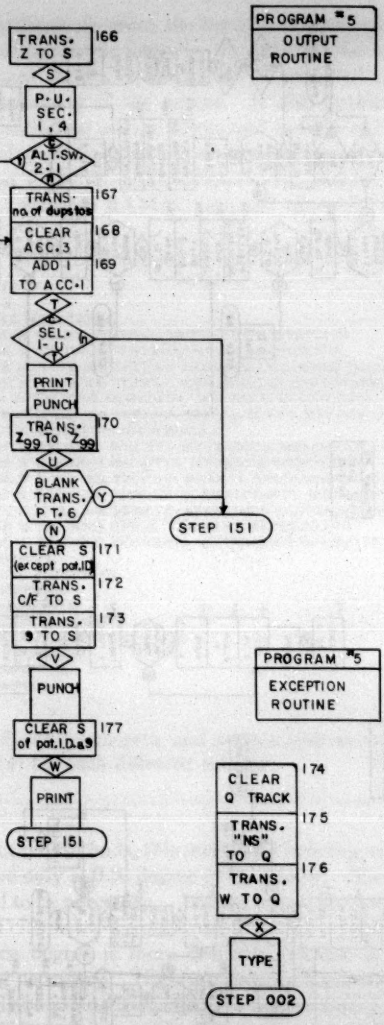


FIG. 5. Flow chart continued.

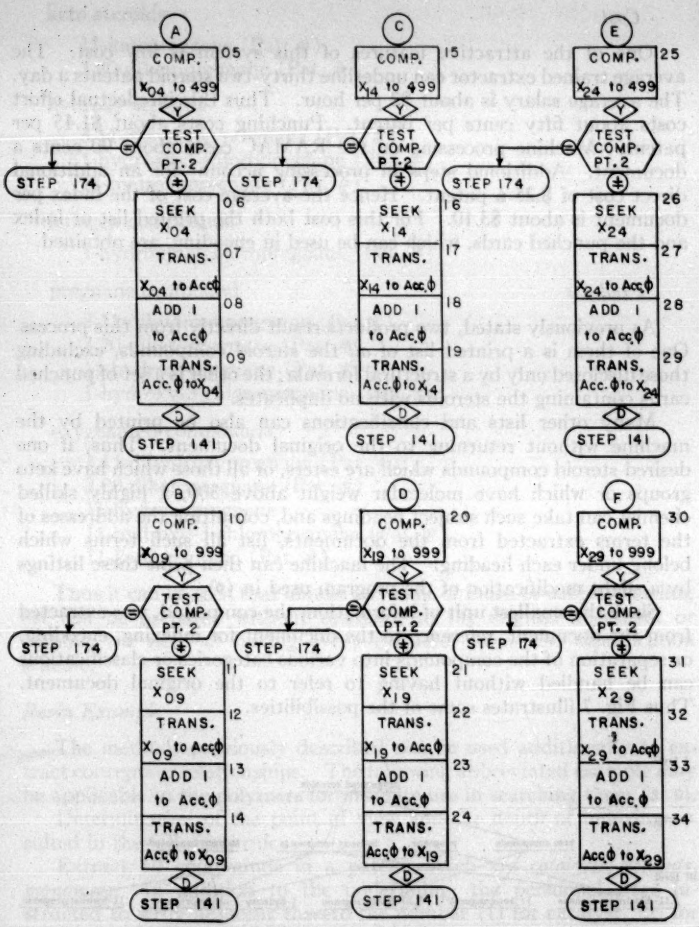


FIG. 6. Flow chart illustrating the letters A, B, C, D, E, and F as examples.

Training Time

Relatively little time is required to train an indexer adequately to recognize steroid compounds, rarely more than a week. His training consists of supervised extraction and each of his errors is discussed in detail with him during the training period.

Cost

One of the attractive features of this system is low cost. The average trained extractor can underline thirty-two steroid patents a day. The average salary is about $2 per hour. Thus this intellectual effort costs about fifty cents per patent. Punching costs about $1.45 per patent. Machine processing on the RAMAC costs about 90 cents a document. Additional steps in processing account for an additional direct cost of $.25 a patent. Hence the average cost of the index per document is about $3.10. For this cost both the printed list or index and the punched cards, which can be used in encoding, are obtained.

Products

As previously stated, two products result directly from this process. One of them is a printed list of *all* the steroid compounds, excluding those disclosed only by a structural formula; the other is a set of punched cards containing the steroids with no duplicates.

Many other lists and classifications can also be printed by the machine without returning to the original document. Thus, if one desired steroid compounds which are esters, or all those which have keto groups or which have molecular weight above 500, a highly skilled chemist can take such subject headings and, consulting the addresses of the terms extracted from the documents, list all such terms which belong under each heading. The machine can then print these listings by a slight modification of the program used in (9).

Since the smallest unit of information, the compound, was extracted from the document, reference to the document for indexing, encoding, or separation of the compounds into various categories or classifications can be handled without having to refer to the original document. Thus Fig. 7 illustrates some of the possibilities.

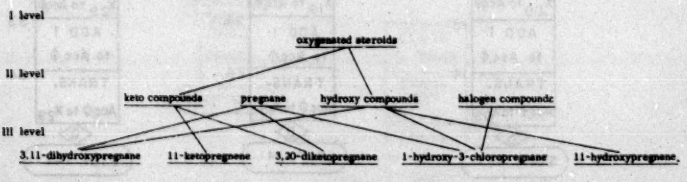


FIG. 7.

Level III is essentially the original index. Levels I and II can be created in a manner similar to the subject matter headings by one skilled person. It is now easy to prepare other lists such as:

keto steroids

11-ketopregnene (Pat. #)
3,20-diketopregnane (Pat. #)

hydroxy steroids

3,11-dihydroxypregnane (Pat. #)
1-hydroxy-3-chloropregnane (Pat. #)
11-hydroxypregnene (Pat. #)

halogen compounds

1-hydroxy-3-chloropregnane (Pat. #)

pregnane compound

3,11-dihydroxypregnane (Pat #)
3,20-diketopregnane (Pat. #)
11-hydroxypregnane (Pat. #)
1-hydroxy-3-chloropregnane (Pat. #)

oxygenated compounds

3,11-dihydroxypregnane (Pat. #)
3,20-diketopregnane (Pat. #)
11-hydroxypregnene
11-hydroxy-3-chloropregnane (Pat. #)
11-ketopregnane (Pat. #)

Thus it can be seen that almost any type of index or subject heading list can be prepared without reference to the original document or without the need of having every extractor know the classification scheme.

Resin Example

The methods previously described can be used additionally to extract concepts or relationships. The following abbreviated example may be applicable to the polymers for ultimate use in searching them (**3, 9**).

Determination of the point of view and the depth of indexing resulted in the following rule:

Extract *all* compounds in a patent which are *catalysts*, *solvents*, *monomers*. In addition to the underlining, the personnel were instructed to write adjacent thereto the number (1) for catalyst, (2) for solvent and (3) for monomer. If the number of kinds of different concepts or relationships is kept small, the extractors work satisfactorily. A specific example is shown in Tables III, IV, and V. Indexes can easily be printed for all catalysts, all solvents, all monomers, or all compounds in the patent regardless of their function. (See Fig. 8.)

Patented Nov. 16, 1943 2,334,195

UNITED STATES PATENT OFFICE

2,334,195

PROCESS FOR THE PRODUCTION OF POLYMERIZATION PRODUCTS

Heinrich Hopff and Siebert Goebel, Ludwigshafen-on-the-Rhine, Germany, assignors, by mesne assignments, to General Aniline & Film Corporation, New York, N. Y., a corporation of Delaware

No Drawing. Application March 6, 1940, Serial No. 322,547. In Germany March 14, 1939

5 Claims. (Cl. 260—80)

The present invention relates to a process for the production of polymerization products.

The copending application Ser. No. 309,238, filed December 14, 1939, by the present inventors and Curt W. Rautentrauch, describes a process for the production of polymerization products in which ethylene if desired in admixture with other polymerizable compounds, is subjected to polymerization in an aqueous emulsion in the presence of substances supplying oxygen.

We have now found that ethylene may be converted into solid or semi-solid polymerization products in a very simple manner by effecting its polymerization in the presence of organic solvents and oxygen or substances supplying oxygen. The organic solvents may be used either alone or in combination with aqueous liquids in which the solvents and the ethylene are emulsified or dissolved. By our new method of working the solid polymerization products may be obtained under relatively low superatmospheric pressures. Comparatively small proportions of organic solvents, say from 10 to 100 per cent (percentage by weight with reference to ethylene) are sufficient for obtaining good yields of solid polymerization products at between about 100 and about 300 atmospheres, whereas considerably higher pressures are required when conducting the process in the absence of organic solvents.

Among the substances having an accelerating effect in the polymerization we may mention oxygen or substances supplying oxygen, especially peroxidic substances for example persulphates, percarbonates, perborates as well as benzoylperoxide, peracetic acid, diacetylperoxide, toluic acid peroxide and oleic peroxide which are obtainable from the acid chlorides by means of hydrogen peroxide and caustic soda solution.

As suitable organic solvents we may mention by way of example benzene, toluene, xylene and halogenated hydrocarbons, such as ethylene chloride, carbon tetrachloride and chlorbenzene.

Interpolymerization products may be produced by treating ethylene together with further unsaturated substances which in their turn are capable of polymerization under the conditions employed, i. e. substances capable of polymerizing alone under the conditions concerned and substances which polymerize only when exposed to the polymerizing conditions in conjunction with substances polymerizable by themselves. Among substances polymerizable by themselves we may mention by way of example propylene, isobutylene, butadiene, styrene and acrylic acid esters; compounds not polymerizable by themselves but in conjunction with polymerizable substances are for example maleic acid diethyl ester and dimethyl ester. By varying the proportions of the components employed, the properties of the polymerization products may be varied considerably.

The polymerization products are suitable for a great variety of applications. They may be worked as plastics by injection-moulding or used for coating electrical conductors or for the manufacture of condensers. When dissolved in organic solvents they may be used for the preparation of coatings resistant to acids and alkalies.

The following examples serve to illustrate how our present invention may be carried out in practice, but the invention is not restricted to the said examples. The parts are by weight.

Example 1

100 parts of liquid ethylene, 100 parts of benzene and 2 parts of benzoyl peroxide are heated in a pressure vessel at from 80 to 100° C. After a short time the initial pressure of 100 atmospheres falls to about 70 atmospheres. After removing any unpolymerized ethylene and distilling off the solvent, the polymerization product is left behind in the form of a white waxy mass melting above 100° which may be purified by redissolving from organic solvents.

When using toluene instead of benzene a valuable polymerization product is likewise obtained.

If trichlorethylene is employed instead of benzene an interpolymerization product containing chlorine is obtained having similar properties.

Example 2

250 parts of benzene and 500 parts of ethylene are emulsified in a pressure vessel in 1000 parts of an aqueous solution containing 20 parts of the sodium salt of alpha-hydroxyoctodecane sulphonic acid, 1.5 parts of potassium persulphate, 7.5 parts of hydrogen peroxide and 10 parts of benzoyl peroxide and polymerized under a pressure of 150 atmospheres at between 100° and 110° C. After 2 hours the emulsion is precipitated by the addition of an electrolyte, such as common salt. A white powder is obtained having a melting point of above 100° C.

Example 3

100 parts of ethylene, 200 parts of methanol and 2 parts of benzoylperoxide are mixed in a pressure resisting vessel provided with a stirring device and heated to from 110° to 120° C. for 2 hours. The pressure, being 80 atmospheres at the outset, rises to 160 atmospheres and falls

FIG. 8. Patent underlined for solvents, monomers and catalysts.

TABLE III.—*A List of All Solvents in Patent No. 2,334,195.*

The two-digit number at extreme right indicates one less than the number of times the compound occurs (also true of Tables IV and V).

2334195	BENZENE-2	06
2334195	BUTANOL-2	01
2334195	CARBON TETRACHLORIDE-2	00
2334195	CHLORBENZENE-2	00
2334195	CHLORINATED ALIPHATIC HYDROCARBON-2	00
2334195	DICHLORETHYLENE-2	00
2334195	ETHYLENE CHLORIDE-2	00
2334195	HALOGENATED HYDROCARBONS-2	00
2334195	METHANOL-2	06
2334195	ORGANIC SOLVENTS-2	05
2334195	ORGANIC SOLVENT-2	01
2334195	SOLVENTS-2	00
2334195	SOLVENT-2	01
2334195	TOLUENE-2	01
2334195	TRICHLORETHYLENE-2	00
2334195	XYLENE-2	00

TABLE IV.—*A List of All the Monomers in Patent No. 2,334,195.*

2334195	ACRYLIC ACID ESTERS-3	00
2334195	ACRYLIC ACID ETHYL ESTER-3	00
2334195	BUTADIENE-3	00
2334195	ETHYLENE-3	19
2334195	ISOBUTYLENE-3	01
2334195	PROPYLENE-3	00
2334195	STYRENE-3	01
2334195	UNSATURATED SUBSTANCES-3	00
2334195	UNSATURATED COMPOUND-3	00

Phosphate Example

There are seven people now extracting information by this method from the 1,765 phosphorus ester patents in Class 260 subclass 461.

The personnel are instructed to underline all phosphorus compounds. This results in having both generic and specific compounds extracted. The resulting list is an index to all the phosphorus compounds in a patent, excluding structural formulae. (See Table II.) A list can be produced containing all the phosphorus compounds in all the 1,765 patents and under each compound listing the patents in which they occur, for example:

Trimethyl phosphate	Tributyl phosphate	Tricresyl phosphate
Pat. No.	Pat. No.	Pat. No.
Pat. No.	Pat. No.	Pat. No.
Pat. No.	Pat. No.	Pat. No.

No attempt is made to eliminate or correlate synonyms since the search system provides means to allow such redundancy. The index will be quite accurate because of the redundancy of extraction, important compounds usually appearing several times in a document.

TABLE V.—*A List of All of the Catalysts in Patent No. 2,334,195.*

The 2-digit number plus one at the extreme right is the number of times the compound occurs in the patent.

2334195	BENZOYL PEROXIDE-1	05
2334195	BENZOYLPEROXIDE-1	00
2334195	DIACETYLPEROXIDE-1	00
2334195	HYDROGEN PEROXIDE-1	01
2334195	OXYGEN-1	03
2334195	OLEIC PEROXIDE-1	01
2334195	PEROXIDIC SUBSTANCES-1	00
2334195	PERSULPHATES-1	00
2334195	PERCARBONATES-1	00
2334195	PERBORATES-1	00
2334195	PERACETIC ACID-1	00
2334195	POTASSIUM PERSULPHATE-1	01
2334195	PEROXIDE-1	00
2334195	SUBSTANCES SUPPLYING OXYGEN-1	04
2334195	TOLUIC ACID PEROXIDE-1	00
2334195	ACRYLIC ACID ESTERS-3	00
2334195	ACRYLIC ACID ETHYL ESTER-3	00
2334195	BUTADIENE-3	00
2334195	ETHYLENE-3	19
2334195	ISOBUTYLENE-3	01
2334195	PROPYLENE-3	00
2334195	STYRENE-3	01
2334195	UNSATURATED SUBSTANCES-3	00
2334195	UNSATURATED COMPOUND-3	00
2334195	BENZENE-2	06
2334195	BUTANOL-2	01
2334195	CARBON TETRACHLORIDE-2	00
2334195	CHLORBENZENE-2	00
2334195	CHLORINATED ALIPHATIC HYDROCARBON-2	00
2334195	DICHLORETHYLENE-2	00
2334195	ETHYLENE CHLORIDE-2	00
2334195	HALOGENATED HYDROCARBONS-2	00
2334195	METHANOL-2	06
2334195	ORGANIC SOLVENTS-2	05
2334195	ORGANIC SOLVENT-2	01
2334195	SOLVENTS-2	00
2334195	SOLVENT-2	01
2334195	TOLUENE-2	01
2334195	TRICHLORETHYLENE-2	00
2334195	XYLENE-2	00

AUTOMATIC ENCODING

From the punched cards so created the document can be automatically encoded. A detailed illustration of the coding of the steroids will now be described. Although not described in this paper, the phosphates and polymers can be semi-automatically or automatically encoded by using similar procedures and techniques.

A slight modification of the first step (extracting) requires that the compound not only be underlined but that a virgule be placed to separate the various selected parts of the compound, for example (see Fig. 9 and Table VI):

3,20-diketo/pregnane/.
3-hydroxy/11-keto/17-methyl/androstane/.
Δ^4-pregnene/3,11,20 trione/.

United States Patent Office 2,871,246 Patented Jan. 27, 1959

2,871,246

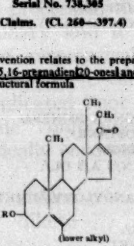
PROCESS FOR 3-HYDROXY-6-ALKYL-5,16-PREGNADIEN-20-ONES AND ESTERS THEREOF

Bjarte Lökes, Roosevelt, Puerto Rico, assignor to G. D. Searle & Co., Chicago, Ill., a corporation of Delaware

No Drawing. Application May 28, 1958 Serial No. 738,305

5 Claims. (Cl. 260—397.4)

The instant invention relates to the preparation of 3β-hydroxy-6-alkyl-5,16-pregnadien-20-ones and their esters of the general structural formula



In the foregoing structural formula the lower alkyl radical in the 6-position can be a straight-chain or branched propyl, butyl, amyl or hexyl group, but it is preferably a methyl or ethyl radical. The radical R can represent hydrogen or an acyl radical of a hydrocarbon carboxylic acid, preferably of less than 7 carbon atoms. Examples of such acyl groups are formyl, acetyl, propionyl, butyryl, valeryl and hexanoyl.

The foregoing compounds are progestationally active. They also serve as starting materials for the preparation of 6-alkylprogesterones and 6-alkyl-17α-hydroxyprogesterones and their 17-esters which are potent progestational agents.

The procedures in the literature for conversion of diosgenin or cholesterol to pregnadiene derivatives all suffer from serious disadvantages in specific steps of the reaction sequence. Actually only the first conventional step of isomerizing a 6-alkyldiosgenin to its corresponding pseudoderivative, 3β,26-diacetoxy-6-alkyl-5,20(22)-furostadiene can be conducted by a procedure analogous to that generally employed for diosgenin, namely reaction with acetic anhydride for 6 hours at 190–195° C. The subsequent steps of the sequence had to be altered materially and in some steps, as in cleavage of the 3-acetoxy-γ-methylpentanoyl ester, entirely different procedures had to be devised.

Thus, the second step, namely oxidation of the furostadiene to the 16-(3-acetoxy-γ-methyl)pentanoate by chromium trioxide in an acetic acid medium is advantageously effected in the presence of a chloralkane solvent amounting to 15–60% by volume of the oxidation medium. The chlorinated solvents contemplated are the chlorinated hydrocarbons of one to three carbon atoms including, for example, ethylene chloride, chloroform, carbon tetrachloride, ethylene trichloride and propylene chloride. Effecting the oxidation in this manner facilitates recovery of the steroid from the reaction mixture. Thus the chromium salts and acetic acid are readily removed in aqueous phase while the 3-acetoxy-16-(3-acetoxy-γ-methylpentanoyloxy)-5-pregnen-20-one stays in the chlorinated hydrocarbon organic phase; this ester can then be recovered by evaporating the chlorinated hydrocarbon.

The subsequent cleavage is effected by a novel solvolysis with at least three equivalents of alkali in aqueous acetone. A rapid reaction occurs with very strong alkali, e. g. caustic soda and potash. More time is needed with weaker bases such as the alkali metal carbonates. The acetone-water ratio can vary from about 1:3 to 3:1. The alkali, e. g. sodium hydroxide, can constitute from 1–20% by weight of the alkaline medium. The time period for the treatment is relatively short, ranging from about 5–60 minutes. Treatment temperatures are in the range of 20° C. to about 65° C. Temperature is relatively critical because two different results are attainable. In the lower range of about 20–33° C., cleavage of the substituted pentanoate results in a high yield of 3β-acyloxy-6-alkyl-5,16-pregnadien-20-one. On the other hand, in the range of about 45–65° C. there also occurs deacetylation of the 3-acylate to yield 3β-hydroxy-6-alkyl-5,16-pregnadien-20-one.

It is within the contemplation of the instant invention to carry out the same alkaline acetone-water treatment further at higher temperature to hydrolyze 3β-acyloxy-6-alkyl-5,16-pregnadien-20-one to 3-hydroxy-6-alkyl-5,16-pregnadien-20-one.

These compounds are valuable as key intermediates in the overall conversion of diosgenin into medically active 6-alkyl steroids. The 3β-hydroxy-6-alkyl-5,16-pregnadien-20-ones have certain advantages over their 3-esters as intermediates. Epoxidation to 3β-hydroxy-6-alkyl-16α,17α-epoxy-5-pregnen-20-ones can be effected by reaction with alkaline hydrogen peroxide in methanol solution to give a better yield. Also the 3-hydroxy compounds can be converted by the Oppenauer oxidation to the 6-alkyl-5,16-pregnadien-3,20-diones and then by Raney nickel reduction to progestationally active 6-alkylprogesterones or epoxidation as above outlined to the 16α,17α-epoxides, followed by conversion to the 17α-hydroxy compounds according to conventional practices. The resulting esters of 17-hydroxyprogesterones are highly potent progestational agents. It has also been found that the 3-hydroxy-6-alkyl-5,16-pregnadien-20-ones are progestational agents.

The following examples describe the invention in further detail but are not to be construed as limiting in spirit or in scope. Quantities are given in parts by weight.

Example 1

To an ice-cooled solution of 100 parts of diosgenin in 1340 parts of dichloromethane there is added a solution of 160 parts of acetone, 270 parts of dichloromethane, 2 parts of sodium acetate and 75 parts of peroxyacetic acid. After completion of the addition, the mixture is left at room temperature for 5 hours and then washed with 1000 parts of a 5% ferrous sulfate solution which serves to reduce the excess peroxyacetic acid. The organic layer is separated and carefully washed with water. The washings are reextracted with dichloromethane and the combined dichloromethane extracts are taken to dryness on a steam bath. Then 900 parts of heptane are added and a small head fraction is distilled off in order to remove residual dichloromethane. The mixture is cooled to 30° C. under agitation and the crystalline material is collected on a filter. This material consists of a mixture of the α- and β-epoxides; the optical rotation α_D of −120° in chloroform solution indicates predominance of the α-epoxide. The mother liquor is evaporated to a small volume and a second crop is harvested with a rotation α_D of −107°, which is also a mixture of the two epimeric epoxides, but in this crop the β-epoxide predominates. The first crop is dissolved

FIG. 9. Steroid patent underlined for automatic encoding.

The duplicates are eliminated from the punched cards as previously described. Before automatically encoding, a dictionary must be generated.

TABLE VI.—*List of Compounds Extracted from Underlined Patent (Fig. 9).*

The two-digit number is one less than the number of times the compound occurs in patent.

2871246	6 ALKYL/PROGESTERONES/.	01
2871246	6 ALKYL/17A/17HYDROXY/PROGESTERONES/.	00
2871246	6 ALKYL/DIOSGENIN/.	00
2871246	3 ACYLOXY/16-D ACETOXY G METHYLPENTANOYLOXY-/5 PREGNEN/20ONE/.	01
2871246	6 ALKYL/5 16PREGNADIEN/3 20DIONES/.	00
2871246	6 A/6 METHYL/16A/17A/1617EPOXY/4 PREGNENE/3 20DIONE/.	00
2871246	6 A/6 METHYL/17A/17HYDROXY/PROGESTERONE/.	00
2871246	3 ACVLATES/3 B/3 HYDROXY/6 METHYL/5 16PREGNADIEN/20ONES/.	00
2871246	3 ACYLOXY/26ACETOXY/6 METHYL/5 20-22-FUROSTADIENE/.	00
2871246	3 ACETOXY/6 METHYL/16-D ACETOXY G METHYLPENTANOYLOXY-/5 PREGNEN/20ONES	00
2871246	./.	00
2871246	3 B/3 HYDROXY/6 ALKYL/5 16PREGNADIEN/20ONES/.	01
2871246	3 B/3 26DIACETOXY/6 ALKYL/5 20-22-FUROSTADIENE/.	00
2871246	3 B/3 ACVLOXY/6 ALKYL/5 16PREGNADIEN/20ONE/.	01
2871246	3 B/3 HYDROXY/6 ALKYL/5 16PREGNADIEN/20ONE/.	00
2871246	3 B/3 HYDROXY/6 ALKYL/16A/17A/1617EPOXY/5 PREGNEN/20ONES/.	00
2871246	3 B/3 HYDROXY/5 A/6 A/5 6 EPOXY/22A/22ALLOSPIROSTANE/.	00
2871246	6 B/6 METHYL/22A/22ALLOSPIROSTANE/3 B/5 A/3 5 DIOL/.	02
2871246	3 B/3 ACETOXY/6 B/6 METHYL/22A/22ALLOSPIROSTAN/5 A/5 OL/.	03
2871246	6 B/6 METHYL/22A/22SPIROSTANE/3 B/5 A/3 5 DIOL/.	01
2871246	3 B/3 ACETOXY/16B/16-D ACETOXY G METHYLPENTANOYLOXY-/6 METHYL/5 PREGNE	01
2871246	.N/20ONE/.	01
2871246	3 B/3 HYDROXY/6 METHYL/5 16PREGNADIEN/20ONE/.	05
2871246	3 B/3 ACETOXY/5 A/6 A/5 6 EPOXY/22A/22SPRIOSTANE/.	00
2871246	6 B/6 ETHYL/22A/22ALLOSPIROSTANE/3 B/5 A/3 5 DIOL/.	00

Dictionary Generation

Two methods are available for generating a dictionary. In the first method one attempts to list all the possible entries. There are two serious defects to this method, (1) it is impractical and time-consuming to think of all the entries, and (2) some of the entries never used must still be stored at no small cost. The second method is to use the extracts from the documents in a suitable computer program to generate the dictionary. This was experimentally and successfully tried with the steroids. Extracts from the first document were put into the machine, for example,

3,20-diketo/pregnane/.
3-hydroxy/pregnane/.
11-hydroxy/pregnene/.

From these, the program generated the following terms which were alphabetized and listed to form a dictionary:

diketo	pregnane
hydroxy	pregnene

The extracts from the next patent contained the following compounds:

3-hydroxy/pregnene/.
3,20-diketo/pregnane/.
11-hydroxy/pregnane/.
3-hydroxy/7-chloro/pregnane/.

The first three compounds were automatically processed since the dictionary already contained the terms:

diketo	pregnane
hydroxy	pregnene

and these compounds contain only dictionary terms even though different from the compounds in the first patent. However, when the fourth compound is processed, the machine recognizes that there is no *chloro* in the dictionary and it prints out on the typewriter "*Chloro—not in File,*" and automatically punches out a card which contains the whole compound. At the end of the run, all the terms which the typewriter indicated "not in file" are added to the dictionary. Now the "not in file" compounds which were not processed are sent through and automatically encoded. The new dictionary will now contain

chloro	hydroxy	pregnene
diketo	pregnane	

Starting with ten steroid terms, we have generated a dictionary of about seven hundred terms. A small sample of the dictionary is shown in Table VII. The dictionary terms increase about four terms per steroid patent. The average steroid patent contains about fifty compounds. It is hoped that the number of terms added will materially decrease as the size of the dictionary increases. Trivial names such as "progesterone," if listed, are, of course, added to the dictionary. The dictionary is therefore self generating and every term therein has been used at least once. (See Table VII.)

TABLE VII.—*A Partial List of the Steroid Dictionary Generated from the Document Extracts.*

ACETATE	EPOXY	METHYL	TRIONE
ACETOXY	ETHYL	METHOXY	TRIHYDROXY
	FLUORO		TESTOSTERONE
BUTYL	FUROSTADIEN	ONE	TRIOL
BENZOATE		OL	TRIKETO
	HYDROXY	OXIDO	TRIFLUOROACETYL
CAPROATE	HYDROCORTISONE		TETRAHYDROXY
CHOLESTADIENE		PREGNADIENE	TETROL
	ISOPROPYL	PROGESTERONE	
DIONE	ISOBUTYL	PREGNANE	
DIHYDROXY		PREGNATRIENE	VALERYL
DIKETO	KETAL		
DIKETAL	KENDALLS COMPOUND A	SPIROSTEN	
DIESTER	LOWER ALKYL	SITOSTEROL	YAMOGENIN

Encoding

The punched cards generated by the dictionary are fed into the machine. For example, for

3,20-diketo/11-hydroxy/pregnene/.

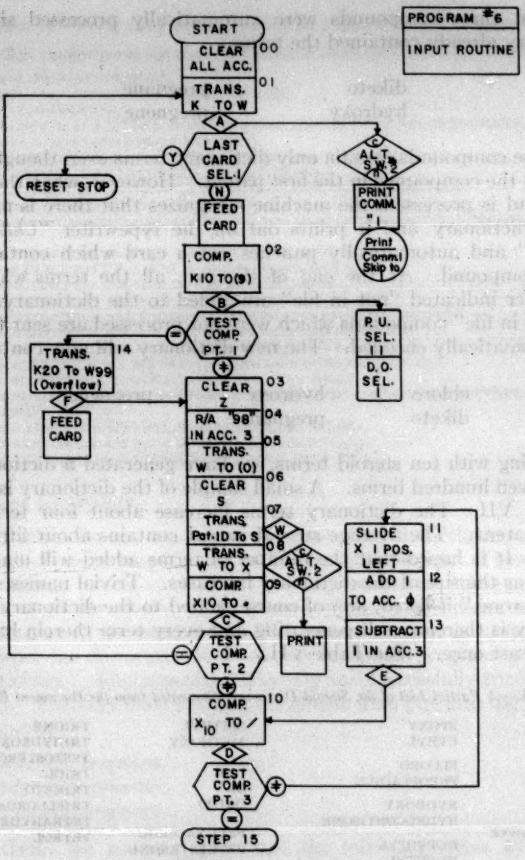


FIG. 10. A flow chart of a program for automatic encoding. This chart is continued in Figs. 11 through 16.

the machine recognizes the first portion indicated by the / and looks up "diketo" in the dictionary. In the dictionary the codes for diketo are stored with the term, provisions for handling the numbers (for example, 3, 20) having already been made. The codes are then sent to the output of the computer. The machine then recognizes and processes "11

hydroxy" and finally "pregnene." The period "." indicates the end of the compound. The codes which have been sent to output are then either printed if there is no suitable punch available or, if one is available, the codes are punched in a punched card. It does not matter how a compound is named in the document. It will always receive the same code. For example:

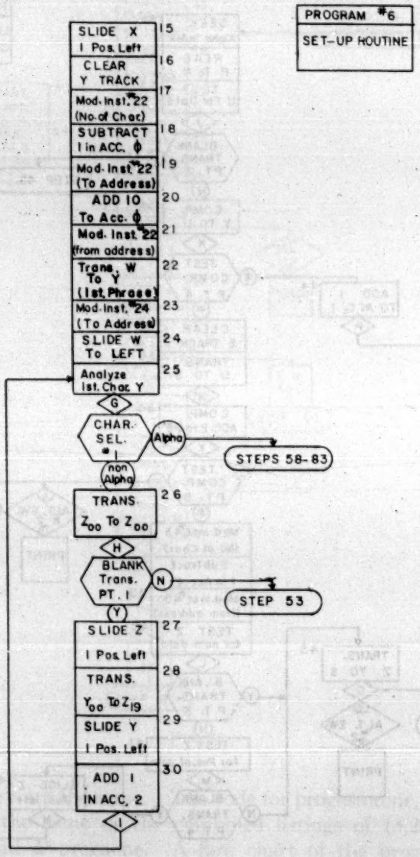


FIG. 11.

3,20-diketo/Δ⁴-pregnene/.
Δ⁴-pregnene/3,20-dione/.
progesterone/.

are all different methods of expressing the same compound, and will each receive the same codes. The machine in the first instance processes

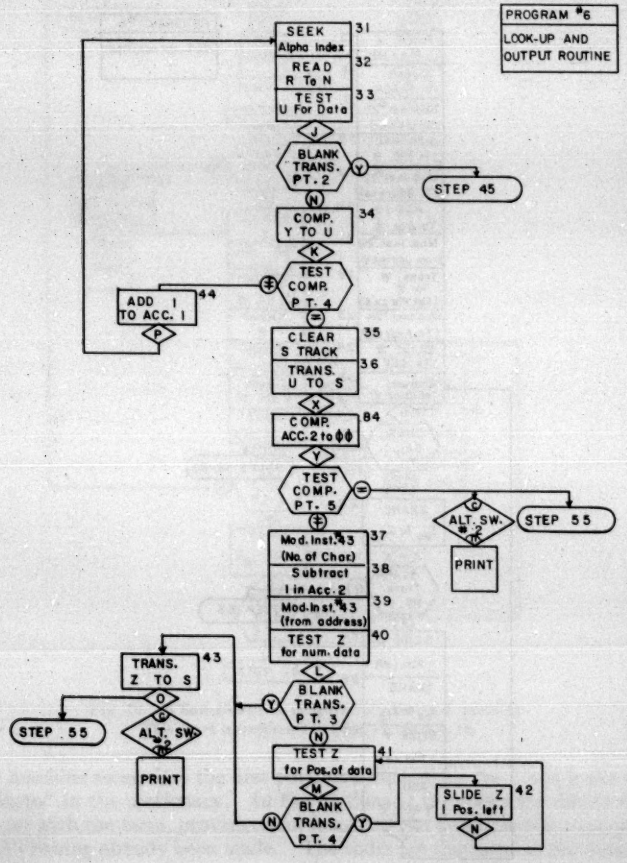


FIG. 12.

diketo and then pregnene; in the second instance it processes pregnene and dione; and, in the third it merely puts out the code for progesterone. The dictionary can thus contain both easily recognized parts of compounds as well as their trivial names. The code for dione and diketo

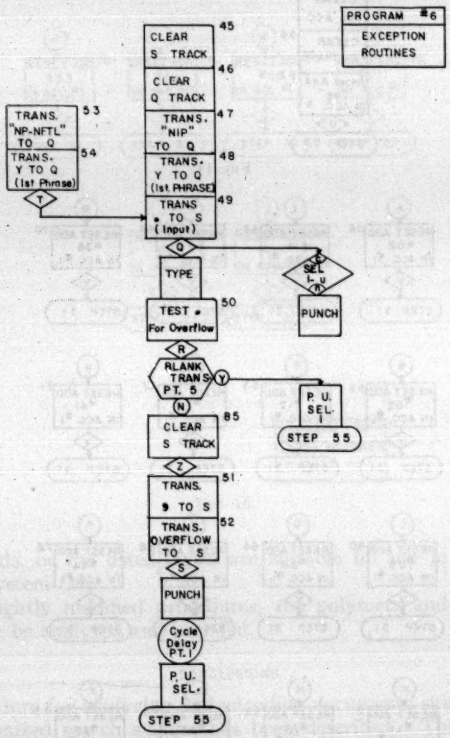


FIG. 13.

in the dictionary are the same. The code for progesterone, listed in the dictionary, is the same as the combined listings of (3,20-diketo) or (3,20-dione) and Δ^4-pregnene. A flow chart of the program for encoding is shown in Figs. 10 through 16.

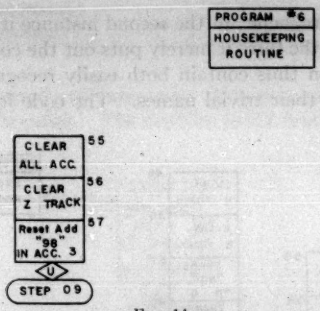

PROGRAM *6
HOUSEKEEPING ROUTINE
CLEAR ALL ACC.
55
CLEAR Z TRACK
56
Reset Add "00" IN ACC. 3
57
U
STEP 09

FIG. 14.

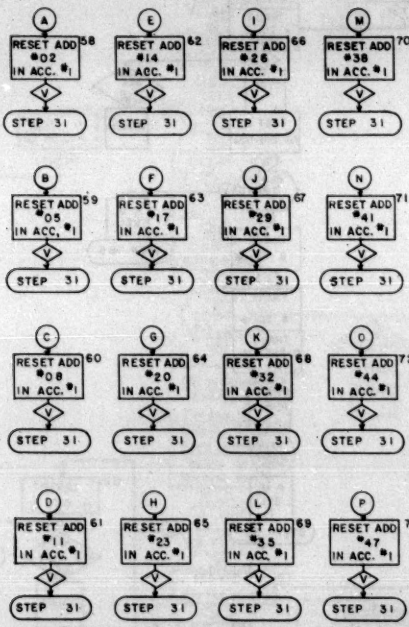

A
RESET ADD *02 IN ACC. *1
58
V
STEP 31
E
RESET ADD *14 IN ACC. *1
62
V
STEP 31
I
RESET ADD *26 IN ACC. *1
66
V
STEP 31
M
RESET ADD *38 IN ACC. *1
70
V
STEP 31
B
RESET ADD *05 IN ACC. *1
59
V
STEP 31
F
RESET ADD *17 IN ACC. *1
63
V
STEP 31
J
RESET ADD *29 IN ACC. *1
67
V
STEP 31
N
RESET ADD *41 IN ACC. *1
71
V
STEP 31
C
RESET ADD *08 IN ACC. *1
60
V
STEP 31
G
RESET ADD *20 IN ACC. *1
64
V
STEP 31
K
RESET ADD *32 IN ACC. *1
68
V
STEP 31
O
RESET ADD *44 IN ACC. *1
72
V
STEP 31
D
RESET ADD *11 IN ACC. *1
61
V
STEP 31
H
RESET ADD *23 IN ACC. *1
65
V
STEP 31
L
RESET ADD *35 IN ACC. *1
69
V
STEP 31
P
RESET ADD *47 IN ACC. *1
73
V
STEP 31
PROGRAM *6
ALPHA ROUTINES

FIG. 15.

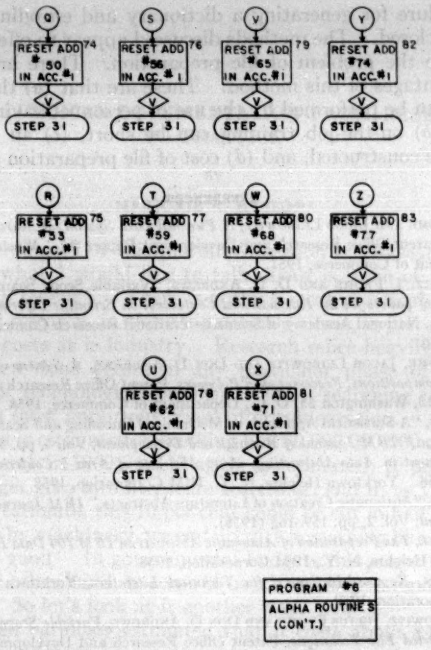


FIG. 16.

The cards, or the listed codes are suitable for use in the steroid searching system (1).

With slightly modified procedures, the polymers and phosphates can likewise be analyzed and encoded.

CONCLUSIONS

A procedure for analyzing and encoding documents suitable for use in a mechanized search system has been described. Three types of indices are prepared:

1. Index of terms, in the language of the document, alphabetically listed for each document
2. Index of terms, in the language of the documents, of all the documents in the file, alphabetically arranged
3. Index of terms under preassigned headings

A procedure for generating a dictionary and encoding compounds has been disclosed. The methods discussed appear to offer a promising resolution to the problem of file preparation. There are several apparent advantages of this method. These are that (*a*) the bulk of the extraction can be performed by the use of personnel having less formal education, (*b*) on-the-job training can be short, (*c*) an accurate and reliable file is constructed, and (*d*) cost of file preparation is reasonable.

REFERENCES

(1) JULIUS FROME AND JACOB LEIBOWITZ, *A Punched Card System for Searching Steroid Compounds*, Patent Office Research and Development Report #7. Washington 25, D. C., Department of Commerce, 1957.

(2) J. LEIBOWITZ, J. FROME AND D. D. ANDREWS, "Variable Scope Search System: VS_2." *Preprints of papers for the International Conference on Scientific Information*. Washington, D. C., National Academy of Sciences—National Research Council, 1958. Area V, pp. 291–316.

(3) JULIUS FROME, JACOB LEIBOWITZ AND DON D. ANDREWS, *A System of Retrieval—Compounds, Compositions, Processes, and Polymers*, Patent Office Research and Development Report #13, Washington 25, D. C., Department of Commerce, 1958.

(4) H. P. LUHN, "A Statistical Approach to Mechanized Encoding and Searching of Literary Information," *IBM Journal of Research and Development*, Vol. 4, pp. 309–317 (1957).

(5) *An Experiment in Auto-Abstracting—Auto Abstracts of Area 5 Conference Papers*, Nov. 16–21, 1958. Yorktown Heights, N. Y., IBM Corporation, 1958.

(6) H. P. LUHN, "Automatic Creation of Literature Abstracts," *IBM Journal of Research and Development*, Vol. 2, pp. 159–165 (1958).

(7) T. R. SAVAGE, *The Preparation of Automatic Abstracts on IBM 704 Data Processing System*, Yorktown Heights, N. Y., IBM Corporation, 1958.

(8) H. P. LUHN, *Keyword-in-Context for Technical Literature*, Yorktown Heights, N. Y., IBM Corporation, 1959.

(9) JACOB LEIBOWITZ, JULIUS FROME AND DON D. ANDREWS, *Variable Scope Patent Searching by an Inverted File Technique*, Patent Office Research and Development Report #14, Washington 25, D. C., Department of Commerce, 1958.

(10) C. K. SCHULTZ AND J. J. O'CONNOR, "Designing More Efficient Indexes," submitted for publication *UNESCO Bulletin*, Nov.-Dec., 1959.

(11) J. T. MACMILLAN AND I. D. WELT, "Some Personnel Problems of a Small Indexing Project," delivered at ACS 136th meeting in Atlantic City, Sept. 14, 1959. (Unpublished.)

(12) I. D. WELT, "A Combined Indexing-Abstracting System," *Preprints of Papers for International Conference on Scientific Information*. Washington, D. C., National Academy of Sciences—National Research Council, 1958, Area 2, pp. 135–145.

(13) S. M. NEWMAN, *Linguistic Problems in Mechanization of Patent Searching*, Patent Office Research and Development Report #9, Washington 25, D. C., Department of Commerce, 1957.

(14) H. P. LUHN, *Review of Information Retrieval Methods*, Research Report #RC-59. Yorktown Heights, N. Y., IBM Corporation, Oct. 1, 1958.

V.2
196

METALLURGICAL DOCUMENTATION FROM RESEARCH TO PRACTICE

BY

MARJORIE R. HYSLOP[1]

While the subject of this morning's session is "Research in Documentation," what I would like to talk about is some of the practical results of such research. There is nothing more practical than money, and certainly research, in these inflationary times, is just as concerned with cutting costs as is industry. Research relies heavily on knowing what has been done before in any particular phase of a scientific, engineering or technological problem. Thus, searching the published literature is a most important phase of documentation.

Now, if, as is estimated, something like $9,000,000,000 a year will be spent on research in 1960; and if, again as estimated, 10 per cent of this research budget goes into literature searching; and if, as I am sure is a conservative estimate, this 10 per cent could be cut in half by searching the literature by machinery, we end up with something like $500,000,000 a year to the good! In governmental budgets this may not sound like much; on the other hand, to some of us it's too *large* a figure to be comprehensible. So let's look at it another way.

It has been variously estimated that it costs about $40,000 a year including all expenses to maintain one investigator in a laboratory. This figures roughly to $20 per hour. If this researcher on tackling a new assignment spends two weeks in the library (and this is very much on the low side of the time that a search requires) $1600 is gone before he has even started his job, and chances are that he still hasn't found much of the information that actually does exist. If he hasn't found the information, then he must take the additional time and money to perform experiments that may already have been done.

"Well," you say, "I rely on my librarian to make such searches." But librarians are not cheap either. A number of large technical libraries offer literature searches on demand and the going rate is something like $7.50 per hour, plus the cost of whatever clerical and typing labor is involved. This totals less than half the price of the researcher's time, but even so it is still too expensive.

As recognition of the problem grew, it was natural that the idea of doing this laborious repetitive work by machine should emerge. In business and industry the efficiency of machines for keeping records

[1] Managing Editor, *Metal Progress;* and Editor, *ASM Review of Metal Literature,* American Society for Metals; Novelty, Ohio.

and sorting information is well known; in recent years the engineer has grown to rely on machines for marvelous mathematical and engineering computations. But the problem of feeding ideas and concepts into machines, the problem of devising electronic circuits that would understand the everchanging language of science and technology required a quite different approach to machine methods.

Work in this field has been going on for something like ten years. Technical people in various scientific disciplines became interested, and of course the one I want to talk about today is metals and metallurgy. The American Society for Metals has been gathering, editing, publishing, circulating and disseminating valuable technical information for more than forty years. Five years ago it was decided that it was high time something was done about improved methods for finding this valuable information when it was needed. This decision led to sponsorship of a research project at the Documentation Center at Western Reserve University. Goals of the research were simple—merely to determine whether or not searching of metallurgical literature by machine was feasible. The research took the form of a pilot-plant operation which, after four years, had proved to the satisfaction of the management of the American Society for Metals that the answer to the question was "Yes." During 1959, therefore, efforts were concentrated on the transition from pilot plant to operation, and, starting in January 1960, the ASM Documentation Service was offered to the public. This, we believe, is the first practical application of machine literature searching to be offered on a wide-scale operating basis.

V.2
196

Before describing the types of services subscribers will receive, it might be in order to tell something about the input of the machine system—the "machine library." In the extremely broad and wide-ranging field of metals and metallurgy, it is necessary for the American Society for Metals to define its area of interest. This can be done in a brief paragraph. For ASM purposes, "metallurgy means the arts and sciences underlying the production of any metal from its ore or concentrate; its refining, alloying and manufacture into mill shapes including foundry and forge; machining, forming or shaping operations where quality of metal or tool is critical; heat treatment and welding. It includes research into and production of conventional metals and alloys and new variations to meet special services. It includes data on properties, and on performance in any environment. It includes equipment and methods for inspection and research. It includes metallurgical education and history."

This is what we call the "core" of metallurgy. Within the limits of this definition, the American Society for Metals publishes in its abstracting journal, known as the *Review of Metal Literature*, approximately 12,000 abstracts per year taken from the world's published literature. Now we recognize the fact that every metallurgist has

interests that range widely into peripheral fields of science and technology—inorganic chemistry, solid state physics, mechanical engineering, electrical engineering, geology, to name only a few. The ASM also recognized from the beginning that a literature-searching service for the metallurgist, to be completely successful, must either cover these peripheral fields or be of such a nature that it can be coordinated with other literature services covering these fields. A fortunate circumstance has substantially furthered this aim.

This fortunate circumstance is based on the fact that while the feasibility of literature searching by machine has been proven to the satisfaction of the American Society for Metals, there were others more skeptical. One group in particular which feels that machine searching must be proven on a much larger scale than the ASM can possibly afford to underwrite is the National Science Foundation. The philosophy adopted by NSF, however, is exactly that of the ASM, namely, that a large-scale proof of the system requires an operation that will service not only the "core" of metallurgy, but also the peripheral fields where knowledge is buried that the metallurgical researcher must have. The National Science Foundation, therefore, has awarded a grant to Western Reserve University, independent of ASM, to support the machine indexing and abstracting of published literature in the peripheral fields mentioned above—chemistry, physics, mechanical and electrical engineering and a number of others. In other words, ASM pays for encoding into the machine library something like 12,000 documents per year; NSF will supplement this with additional material amounting to 20,000 to 30,000 more abstracts.

At the beginning of the ASM service, which I shall describe shortly, the subscriber will pay only for the cost of running the search (that is, programming and operating the machine and reproducing copies of abstracts) and not for the preliminary costs of compiling the machine library which will be borne jointly by ASM and NSF. It is anticipated, as in the preliminary stage of developing any other new product or process, that costs will be substantially reduced as operating routines are perfected and that before very long the service can become completely self-supporting.

Also what the services are today does not necessarily represent what they will eventually become. As we go into operation we are finding that there are all sorts of possibilities we hadn't even anticipated. Some of these variations of the basic product have been suggested to us by potential subscribers, and we wonder why we didn't think of them ourselves. However, for the time being several basic types of searches are being offered to the beginning subscriber.

The primary version is what we call a "biweekly digest search—special interest." "Biweekly search" means that only current litera-

ture will be surveyed. "Special interest" means that the subscriber specifies exactly the subject area in which he is interested. In other words, the selection of literature is tailored to individual requirements. This service is offered as a subscription at a base price of $50 per month. (It includes searching of the information resources being abstracted under the NSF test program described above.) It provides for the subscriber a selection of abstracts of currently published literature pertinent to his interests, delivered at two-week intervals. The abstracts are typed on separate sheets for filing in any manner he wishes and he can also order photocopies of original documents which are not readily available to him.

A variation of this basic service is the "biweekly digest search—general interest." It is designed to serve more than one subscriber at a somewhat lower cost. In this service a number of subjects will be pre-selected representing problems of current wide interest—for example, vacuum melting and pouring of metals, oxygen processes for steelmaking, welding stainless steels, explosive forming. Since the principal cost of running a machine search is represented by the programming of the machine (that is, wiring it to select appropriate index information), if the same search is sold to more than one subscriber, the cost per subscriber can be materially reduced.

A third variety of documentation service is the bibliographic search delving into past literature. Here the machine work is definitely handicapped at present because the machine library simply does not extend far enough back. Starting with 1960, all of the abstracts published in the ASM *Review of Metal Literature* (some 12,000 per year)[2] are being encoded into machine feed. During 1959, 7500 abstracts representing the most important portion of the 12,000 were fed into the machine. During 1958 the number was 5000; from 1957 to 1955, coverage is spotty and scant, since it represents only that used in the early phases of our research.

To bridge this gap, Information Searching Service will offer bibliographic searching by the time-honored methods using conventional library tools and indexes. Such library work simply cannot be done overnight. It requires much time and skill on the part of a trained reference librarian or literature researcher to ferret out the appropriate index entries in a wide variety of reference sources and then to trail the entries back to the related abstract or title and finally to locate the document itself. But when, in the not too distant future, a complete machine library is available covering a span of ten years—or even five years—it will be no trick at all to program the machine for the

[2] Since last December, when this paper was presented, the NSF test program has become well established, and abstracts are being prepared for machine search at the rate of 25,000 to 30,000 per year covering additional fields of interest and additional types of literature (such as patents) not covered in the published *Review of Metal Literature*.

question, feed in the tapes, and in an hour or less prepare a printed bibliography of all the recent and most important literature.

Another product will be encoded tapes or punched cards of a year's literature for those who want to install a searching machine in their own plants. The advantage to them would be the ability to search any combination or series of desired questions at any time and also to prepare tapes of their own containing information that is either confidential or on subjects not included in the ASM service.

Now a natural question is, "What do I, as a possible subscriber to the ASM Documentation Service, stand to gain? Will it do a better and more thorough job for me than I can get otherwise, will it do it faster, and will it do it cheaper?" We believe that the answer to all three of these questions is an emphatic "Yes."

It would be impossible in the space allotted here to explain the internal mechanics of the machine's "intellectual" operations; it does, however, provide several important functions that partially account for its thoroughness and for its depth of indexing. One is a built-in "code dictionary" which provides synonymous and generic concepts as required, plus some very useful syntax so that the human intellectual labor required to analyze a question is substantially reinforced. It is doubtful that the human factor can ever be eliminated completely, and skilled operators are being trained to make sure not only that the appropriate document analysis is fed into the machine library but also that the questions are properly interpreted in programming the machine for search.

The vastly faster speed of searching by machine is widely known. This results not only from the lightning speed of electronic circuits as contrasted to the slow turning of pages and scanning by eye of index entries, but also from the fact that *all* of the index entries required for a document, whether 10 or 100, can be searched simultaneously. This is something that is certainly beyond the ability of the human brain. On the input side as well as the searching side, increased speed is an important advantage. Why this is so, again, would require an explanation of the machine's internal mechanism which would be out of place here.

We have already said enough about costs and savings that little need be added here. To protect against the danger of the "oversell," however, I should like to point out some limitations of machine searching, because there are certain types of information which can be secured more readily and more cheaply by other means. This can be shown on a graph (Fig. 1) which shows, qualitatively, the cost of machine searching (called "retrieval" by the documentalists) on the vertical ordinate against the type of question or information sought on the horizontal. The solid line represents conventional searching and the dashed line machine searching. The type of question ranges from factual informa-

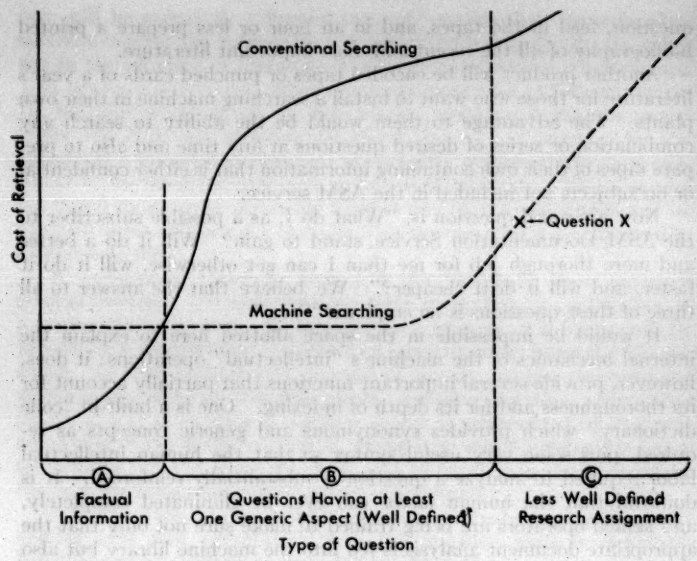


FIG. 1. Cost of retrieval plotted against type of question.

tion to a rather ill-defined research assignment. Following are two examples of requests for factual information (Area A on the graph):[3]

In what proportions or percentages is aluminum soluble in titanium, hafnium and zirconium?

To what temperature and for how long must stainless steel Type 310 be heated to obtain a magnetic transformation?

The curves show that the cost of retrieving factual information is considerably less by conventional methods than by machine searching. In fact the dashed lines for the machine curve in Area A shouldn't be there at all because it isn't possible to retrieve *facts*—at least in the ASM system—but only to locate *sources* where the information might be contained. The conventional searching curve is based on the assumption that the search will be made by a trained librarian or documentalist or scientist who is familiar with the limited number of reference books or similar sources where the data or factual information are likely to be found, and he can therefore go directly to them without much lost time or motion.

[3] Actual requests received in the course of the pilot-plant experiments performed during the research phase of the program.

The central and largest part of the graph is devoted to the questions having at least one generic aspect and couched in terms which are fairly well defined. The cost of searching such questions by machine is approximately the same as the cost of locating factual information by machine (if the latter could be done) because the machine must go through its entire library; it cannot discriminate. However, the cost of searching this type of question by conventional means rises rapidly both because of brain power for analysis of question and labor to scan a large bulk of material. Samples of questions in this area are:

Vacuum melting of nickel-containing materials.
Methods of relieving residual stresses in aluminum alloys.

Finally we come to Area C in the graph which represents questions in which the terminology is vague and difficult to pinpoint from the indexing point of view. A search performed on "notch sensitivity of high-strength steels" is a good example. Even with all the refinements built into the searching machine, both "high-strength steel" and "notch sensitivity" need further definition. Such questions constitute real research problems, and whether the search is done by machine or by conventional tools, some highly skilled library and technical talent is necessary to interpret and define the area of search and analyze the results. The principal reason the cost by machine is considerably lower than by conventional methods is that in the first the operator does only the brain work, in the second the operator also laboriously turns pages.

One final word about costs: Recently we were discussing with a potential subscriber a special and quite complex type of search which would result in generation of an alphabetical subject index to a large group of documents. After exploring the possibilities in a discussion lasting an hour or so, the man finally took a deep breath and asked the inevitable question: "How much is this going to cost?"

His reaction to the figure we quoted surprised even us. It was practically a gasp: "Is this for real? Now I *know* I was born 30 years too soon!"

STORING NUMERICAL VALUES IN AN INVERTED FILE FOR MECHANIZED INFORMATION RETRIEVAL

BY

D. D. ANDREWS [1]

When mechanized information retrieval first started, everyone was content, for the most part, to use qualitative information contained in a document as the basis for machine selecting a document that would have the technical content pertinent to the sought-for information. This qualitative information is, of course, most important and for many purposes may be entirely adequate. However, as we use our present systems of mechanized information retrieval more and more, it is only natural that we seek to improve them and to provide greater flexibility and adaptability for the needs revealed by such actual use.

As an example in point, for more than two years the Patent Office has been mechanically searching the patent literature in the field of steroid chemistry. This search entails essentially a retrieval of all documents which disclose chemical compounds identifiable by their inclusion of certain chemically significant functional groups and the interrelationships among them. This search system has been entirely adequate for the immediate needs of the patent examiners as well as the fifty or so users of this retrieval system outside of the U. S. Patent Office. It was found, however, that very few of the search requests originating in this chemical field are concerned with other than the identification of such chemical compounds either generically or specifically.

When the Patent Office entered the field of polymer resins, a slightly different situation was found. Qualitative information such as the chemical compounds was still important wherever applicable, but, unfortunately, in the field of polymer chemistry it is not possible to identify the polymer product by a precise chemical structure because of the great complexity of these tremendously large molecules. Instead, in the literature dealing with these compounds, the identity of a polymer compound is obtained indirectly by a description of the process used to produce it. A common way to define a polymer compound is to identify not only the monomer used as the starting material in the polymerization process, but also the catalyst, solvent, etc., used during the process. Again, unfortunately, these alone are not, in many cases, sufficiently definitive because entirely different polymer compounds can be produced depending upon the various quantitative or numerical values for the many attendant conditions in the polymerization process. For example, polymerization at 1 or 2 atmospheres of pressure yields a

[1] Director, Office of Research and Development, U. S. Patent Office, Washington, D. C.

wholly different polymer product from that produced if the pressure is in the neighborhood of 100 or more atmospheres. Furthermore, the physical and chemical properties of the polymer product are often relied upon further to specify its composition.

It therefore seemed that steps should be taken to provide the start of a system for mechanically retrieving technical information based, at least in part, upon these quantitative factors of process conditions and properties of the materials resulting from the process. These are here designated as "numerical values." In the field of resins, the numerical values which were selected as having the greatest pertinency were the temperature, pressure, time and pH values of the process and the melting point, molecular weight, specific gravity, crystallinity and tensile strength of the polymer product.

Using these numerical values as a basis for selection in the same way that the qualitative characteristics had been used in the past was found to give very undesirable results. The inherent difficulty was the inability of the person requesting a mechanized search to specify precisely the numerical value or, for that matter, the range of numerical values which he was willing to accept as an answer. Even worse, he could not unequivocally reject all numerical values outside of that range as being not acceptable. No matter what the questioner specified as his minimum and maximum pressures for polymerization, he undoubtedly would want to see or have retrieved for him documents which had all other qualities of his search except that they differed by say a pound or two outside of his specified pressure range. Of course, after he had made a search and found that a great many documents satisfied that search, then he might be willing further to limit the retrieved documents to those having a certain specified set of numerical values. For example, if the technical information being sought concerned polyethylene having a molecular weight in excess of 500,000, it would be quite a hardship to make the search solely on the basis of the monomer used in the polymerization process because hundreds of documents would so respond and the examiner would have a difficult problem in selecting those few which might respond to this numerical value limitation. However, if the numerical value were to be incorporated in the search process, the retrieval might well be restricted to a very few documents.

Additional difficulty arises from the fact that almost invariably numerical values are specified not as a single value but as a range of values. The search request normally is satisfied if a previously described range of numerical values includes, overlaps or touches the numerical value or range of values specified by the search request. As shown in Table I, a number of different situations must be investigated before it can be determined whether a range of numerical values disclosed is or is not pertinent to a given specified range of values.

A further problem in the numerical values selected for the polymers is the greatly differing size of the numbers representing these values; for example, the specific gravity might well be a number considerably less than 1 and the molecular weight might well be a number in excess of 1,000,000. For convenience in handling these numbers in both the minima and maxima of the ranges, it was decided that some form of compression of the numerals was necessary in order to be economical of facilities for storing these values within a search machine. Accordingly, floating decimal point arithmetic was used and an accuracy of three significant figures was considered adequate. The minimum and maximum values accordingly were compressed into ten numerical digits as follows: the first four digits express the minimum value to three significant digits and the power of the multiplier to be used with these three digits. This power may be either positive or negative and thereby secures the maximum range of values within the search system. Similarly, the second four digits are used to express the maximum numerical

TABLE I.—*Numerical Value Ranges.*

	Inquiry		Disclosure	
	Min	Max	Min	Max
Identical inquiry and disclosure	30	60	30	60
Inquiry included within disclosure	35	50	30	60
Disclosure included within inquiry	30	60	35	50
Inquiry and disclosure overlap	30	50	35	60
Inquiry and disclosure "touch"	30	50	50	60
Single value inquiry	35	35	30	50
Single value disclosure	30	60	50	50
"At least" or "More than" inquiry	30	∞	50	60
"Less than" or "Not more than" inquiry	0	35	30	50

value to three significant figures with its power of the multiplier. The last two digits are used to identify the property or condition with which this range is associated and a grouping number within the document itself. Table II shows how the ten-digit code can be used to store and identify a molecular weight ranging from 250,000 to 1,250,000.

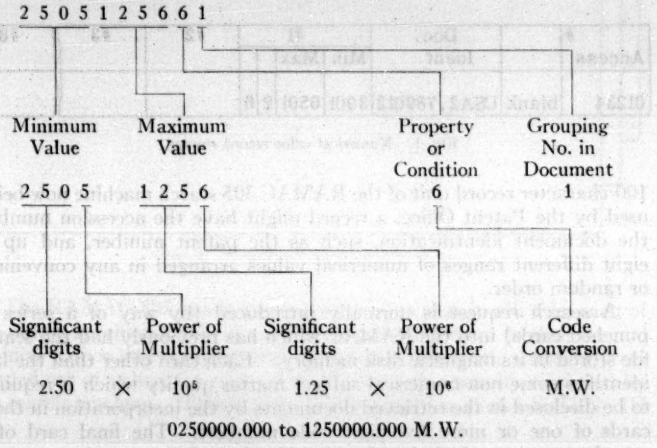
Once it was determined what information was needed and how it was to be economically stored, there was the further problem of how this information was to be processed during a search operation.

As will be pointed out below, the search system devised by the Patent Office for searching polymers is what is sometimes termed an "inverted" file. This term arises from the fact that it is descriptive of the organization of the search file in which is stored information concerning each of the documents included in the search system. Normally all of the information pertaining to a given document is stored in the search file at a single location or a sequence of locations within the main

store and a search is made by scanning the entire store of information in the file to determine which sequence of information units satisfies the search request.

On the other hand, an inverted file takes information units from a single document and stores them throughout the main store at locations predesignated as identifying the subject matter of that information unit. For example, a document assigned an accession number 0001 discloses a process of polymerization wherein ethylene is polymerized by use of a titanium chloride catalyst. In a sequential file, at a location assigned for document No. 0001, codes would be stored signifying

TABLE II.—*Numerical Value Expansion.*
Ten Digit Code



0250000.000 to 1250000.000 M.W.

that it concerned both ethylene and titanium chloride catalyst. On the other hand, in an inverted file at the two locations predesignated for ethylene and titanium chloride catalyst, there would be stored the accession number of that document, 0001. The Patent Office has expanded this inverted file concept so that several levels of genericity can be provided. It is not desirable to attempt to describe this system completely at this time, since this information is already published in Patent Office Research and Development Report No. 14.[2] However, it is evident that the inverted file system makes it very difficult to establish a search term for every possible value or range of values described in the documents included in the search system.

[2] Available from the U. S. Department of Commerce, Washington 25, D. C., for $0.25.

Since, in the search system above referred to, accession numbers for the documents are used only internally of the search machine, an inherent part of the search is the conversion of the accession number to the actual citation data of the documents. This is essentially a "table-look-up" operation in which a part of the storage space of the search machine is reserved for this operation. All of the numerical information pertaining to each document is also stored in the search machine in this reserved portion of the search file. The strategy of the numerical search then resolves itself into making the search of the qualitative information by the usual techniques and superimposing on the results of that search the further information obtained by the comparison of numerical values in the search question and in the stored data. The information stored in the search machine then would look something like Fig. 1. In the

# Access		Doc. Ident.	#1 Min	#1 Max			#2	#3	#8
01234	blank	USA2,789012	3001	6501	2	6			

FIG. 1. Numerical value record storage.

100-character record unit of the RAMAC 305 search machine now being used by the Patent Office, a record might have the accession number, the document identification, such as the patent number, and up to eight different ranges of numerical values arranged in any convenient or random order.

A search request is normally introduced (by way of a series of punched cards) into the RAMAC which has previously had the search file stored in its magnetic disk memory. Each card other than the last identifies some non-numerical subject matter quality which is required to be disclosed in the retrieved documents by the incorporation in these cards of one or more five-digit code numbers. The final card of a question sequence includes the identification of the question or questioner and may either include or not include up to four ten-digit groups denoting four numerical values in which the questioner is interested. Failure to include any numerical values will automatically cause the machine to make a search in the usual fashion and print out the citation data for each document in the search file meeting these non-numerical specifications. However, the inclusion of any one or more of the numerical value ranges in this last or final card of the question set will cause the machine to investigate the information it has in its store concerning the numerical values associated with each document normally retrieved and print out the results of that investigation in a manner to be described presently. Figure 2 shows such a question set in which Examiner J. Doaks requested all documents having ethylene as a

monomer as well as $TiCl_2$ as a catalyst; he was particularly interested in those which mentioned a polymerization process utilizing a pressure between 35 and 65 psi. and those in which the molecular weight of the polymer was between 175,000 and 950,000.

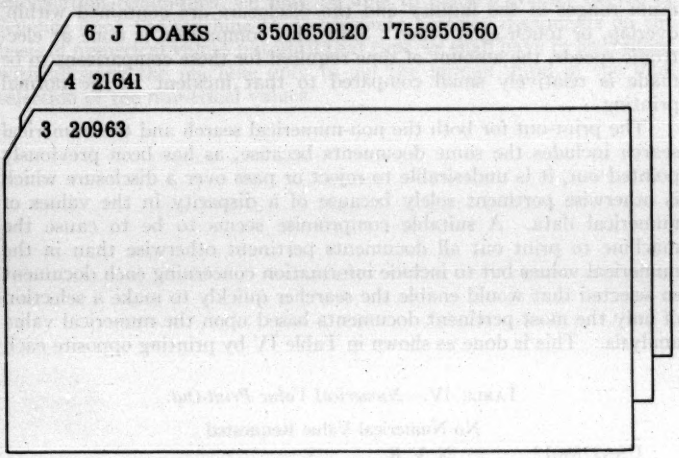


FIG. 2. Numerical value question cards.

The RAMAC 305 has been programmed so that it will take each of the ranges specified by the questioner and expand them to their normal decimal values and compare each expanded value with a similarly expanded set of values for each of the documents normally retrieved by the search machine. This is done rather simply as shown in Table III

TABLE III.—*Numerical Value Comparison.*

	Min	Max
Numerical range of inquiry	35	65
Disclosure in a document	30	50
Reverse Max and Min of inquiry	65	35
Subtract Inquiry from Disclosure	−35	+15

Test signs of differences

Acceptable	0 or −	+ or 0
Not acceptable	+	+, 0, or −
Not acceptable	+, 0, or −	−

in which both the range of the inquiry and the range of the disclosure of a document are subtracted from each other but only after the positions of the maximum and minimum of the inquiry range have been reversed. It is readily seen that the state of the difference—whether it be zero, plus or minus—is a test as to whether or not the maximum and minimum ranges of the inquiry and the disclosure are contained within, overlap, or touch each other. Since this comparison is done at electronic speeds, the amount of time required for these comparisons to be made is relatively small compared to that incident to the normal printing.

The print-out for both the non-numerical search and the numerical search includes the same documents because, as has been previously pointed out, it is undesirable to reject or pass over a disclosure which is otherwise pertinent solely because of a disparity in the values of numerical data. A suitable compromise seems to be to cause the machine to print out all documents pertinent otherwise than in the numerical values but to include information concerning each document so selected that would enable the searcher quickly to make a selection of only the most pertinent documents based upon the numerical value analysis. This is done as shown in Table IV by printing opposite each

TABLE IV.—*Numerical Value Print-Out.*

	No Numerical Value Requested			
USA2789012	N V R			
	Numerical Value Satisfied			
USA2789012	30.000	60.000	K	PAT
	Numerical Value Not Satisfied			
USA2789012	30.000	60.000	K	
	Numerical Value Not Disclosed			
USA2789012	LACK		K	

of the document identification numbers the disclosed range of numerical values for that patent where the inquiry has indicated any interest whatever in that particular numerical value such as pressure or temperature. Additionally, the results of comparing the numerical value ranges of the stored data for each of the selected patents with the range specified in the inquiry are shown by a simple symbol indicating whether the comparison is good or bad. In the example given in Table IV, the letters PAT indicate to the searcher that the range is satisfied. Failure to print such a symbol indicates that the information does not satisfy the search specification but here are the values for the searcher to determine for himself whether they are pertinent. In case a document does

not contain reference to a particular numerical value requested (consequently nothing is stored in the search machine that is pertinent), the word, LACK, is printed in place of any disclosed numerical value range.

Undoubtedly as we become more knowledgeable about mechanized information techniques, we may be able to place a greater degree of reliance upon the machine to discriminate between documents on the basis of numerical values but until that time we shall have to be content with allowing the machine only to aid the human mind in the final selection of the numerical values.

MACHINE LITERATURE SEARCHING IN SCIENCE

BY

ALLEN KENT [1]

INTRODUCTION

For every 60-second period during the day and night—Saturdays and Sundays included—it is estimated that 2000 pages of books, newspapers, or reports will have been published somewhere in the world. If you attempted to keep fully informed of everything going on in the world—through reading—you would fall behind an estimated 1,051,200,000 pages for every solid year you would devote to reading.

This problem faces industrialists, lawyers, military planners, scientists and engineers in keeping abreast of what is published in their own specialized fields. Even in attempting to read the portion published in a single subject field such as chemistry, a chemist would find himself falling behind an estimated 850,000 pages for every solid year he would devote to reading.

PRESENT PRACTICES

In the face of this king-size problem those who need information in order to ensure that they make a reasonable number of well-informed decisions have resorted to three alternate methods for obtaining information they feel they need:

1. They have attempted to be highly selective in their reading, relying on their judgment to select the "one page in a million" which they guess will have the highest probability of containing information of potential future interest.
2. They have attempted to rely on abstracting and indexing services which compress the amount of information that must be scanned.
3. They have attempted to develop special information files for individual or organizational use from special points of view.

These three methods of coping with the increasing literature have their limitations, as follows:

1. Those who try to be highly selective in their reading find themselves losing more and more material because of the unpredictable scatter of material of interest through many publications. Also, if they attempt to float above the flood of information by making their reading general, they find themselves losing some of the gems the reviewers of their field might find inappropriate for inclusion in general discussions.

[1] Associate Director, Center for Documentation and Communication Research, Western Reserve University, Cleveland, Ohio.

Also, the reviewers must themselves cope with the mounting literature and must cut corners in order to get their job done.

2. Those who rely on abstracting and indexing services find themselves facing two problems. First, these services have proliferated tremendously. In a survey I conducted two years ago, I found that more than 300 such services are used by various companies in an attempt to keep up with the literature of their field. Second, the coverage of most of these services is far from complete, and it is difficult to decide which ones to use in order to be confident of reasonably complete coverage of a particular field. In a field such as chemistry, where the American Chemical Society prides itself in relatively complete coverage of the field, they will be preparing more than 130,000 abstracts this year—no mean task in order merely to *scan* the resulting publication.

3. For those who attempt to develop special information files for their own use, it is also necessary to sacrifice completeness of coverage in most cases because of the potential scatter of pertinent information in a tremendous number of publications. The usual budget limitations for this type of work in all but the largest organizations is, of course, an important factor in this regard. Despite these budgetary limitations a considerable amount of money is today being spent on individual information processing activities.

In a questionnaire answered by 604 organizations it was found that 37 per cent of the companies responding prepare abstracts of published papers on a regular basis for the benefit of employees.

It was found that the companies engaging in abstracting and indexing the published literature incurred annual expenses as follows: 40 per cent, more than $5,000; 28 per cent, more than $10,000; 20 per cent, more than $15,000; 7 per cent, more than $35,000; and 4 per cent, more than $50,000.

Then, of course, there are the companies that solve the information problem by essentially ignoring it. An example is the executive of a large steel company insisting it is cheaper to repeat an experiment if the cost does not exceed $100,000 than to pay for a search to determine whether it has been done before.

Why is it so difficult to find information even when it has been indexed consistently and in detail?

Let us consider the problem of exploiting the scientific and technical literature, where the ideas sought are extremely complex and where really difficult correlations are to be made. The problem has three dimensions of frustration:

1. The impossibility of an individual reading and remembering all of the literature that has a reasonable probability of being of later use.

2. The economic impossibility of individuals or their organizations processing for later retrieval the majority of literature of probable pertinent interest.

3. The breakdown of traditional library tools in coping effectively with the detailed requirements of individuals in identifying information pertinent to a given problem.

Consider a question from a jet engine manufacturer, who wants to find a high-temperature-resistant metal possessing the optimum combination of fabricating properties. The inquiry was expressed as a request for all the literature on "the formability of the sheet form of as many of the high-temperature alloys as have been tested." None of the known indexes can handle this type of question conveniently. However, there are so many problems of this sort that are of extreme importance to the American economy and defense effort that there has been a strong reaction in this country aimed at providing specialized information centers—all processing large segments of the literature—many overlapping—from a particular point of view at a particular moment. Here are some of them:

The Titanium Information Center—estimated $1,000,000 a year
Defense Metals Information Center—estimated $1,000,000 a year
Infrared Information Agency
Solid Propellant Information Agency
Prevention of Deterioration Center
Electronic Component Information Center
Chemical-Biological Coordination Center
Cobalt Information Center

Also, there are information centers on the world's literature of: fuel and lubricant additives, nondestructive testing, catalysts, disasters, corrosion, instrumentation, thermophysical properties, deep-sea accidents, chemicals with herbicidal properties, radioisotopes, coffee, bridge hands, stamps.

All this is in addition to the normal indexing and abstracting services that are provided by the various technical societies, such as *Chemical Abstracts*, *Mathematic Abstracts*, *Review of Metal Literature*, *Engineering Index*, *Agronomy Abstracts*, *American Petroleum Institute Technical Abstracts*, *Biological Abstracts*, *Building Science Abstracts*, *Ceramic Abstracts* (U. S. and also England), *Corrosion Abstracts*, *Corrosion et anti-Corrosion Abstracts*, *Corrosion Engineers Abstracts*, *Electrical Engineering Abstracts*, *Electronics Abstracts*, *Fuel Abstracts*, *Iron & Steel Institute Abstracts*, *Packaging Abstracts*, *Rubber Abstracts*, *Resins-Rubbers-Plastics Abstracts*, *Science Abstracts*, *Vacuum Abstracts*, *Vitamin Abstracts*, etc., etc., etc.

I think you will agree that the documentation services of this country are fast becoming out of hand.

V.2
196

POSSIBLE SOLUTIONS

Information Processing System

One apparently ideal solution is the strong centralized information *processing* system of the U.S.S.R.—which receives in one spot the cream of the world's published scientific and technical information and prepares 13 abstract journals and 48 express information bulletins, which criss-cross the 500,000 important published papers with a very minimum amount of overlap and confusion.

The U.S.S.R. system—which undoubtedly has economic advantages—is not, I believe, practical or desirable for this country; we are too far along on our own way. For us, it seems to me, the only reasonable course is the creation of a new type of "library"—a library of encoded information, prepared by the various professional associations, but available for searching, duplication, and distribution anywhere by machine methods, so that centralized *exploitation* as distinguished from centralized *processing* of scientific information can take place. Many separately staffed abstracting services for *processing* and libraries for *exploitation* will operate; but there should be coordination on a national scale of *processing* methods and coverage.

In theory, at least, automation in library procedures makes possible a national system of information facilities that will increase the efficiency of the research worker and safeguard our scientific and industrial progress. But strong disagreement exists among government officials and professional experts in the field as to the degree of centralization desirable and as to the estimated cost of launching any national program. Whenever the subject comes up, wild differences arise in the guesses about costs (one can't call them estimates) which are aired by the various experts talking at the moment. And as always in matters of technology and science these days, someone makes comparisons—in awe or horror—with the Soviet Union, which has a centralized science information processing service employing 2000 staff members.

On the question of costs, I believe that a budget of $36,000,000—to be expended over a period of ten years—would suffice to create an effective national coordination center of scientific information. (I have heard estimates that go up to $300 million.) At the end of this period, I think, these services would become financially self-supporting. In effect, the Center would be handling the new form of "publication"—encoded on magnetic tape or other easily reproducible storage medium—and the returns would amortize the costs of machine searching among a large number of clients.

The existence of a national center would not prevent the formation of other, specialized processing centers, like many publishing houses, whose output might be purchased or rented by organizations wishing to operate their own literature searching machines.

A National Literature Exploitation Center

In order to determine the economic demand for services from a national information exploitation center, a survey was conducted to determine how much various organizations would be willing to pay for comprehensive question-answering services. The 359 organizations that provided usable data indicated an annual economic potential of $2,927,039,153 if reliable exploitation services could be provided.

However, in order to implement such a service, four types of problems must be solved: (1) technical; (2) administrative; (3) cooperative; and (4) political.

Technical Problems

The technical problems involved may be considered from the point of view of what type of service should be provided from a national center. The "punch line" of a system for information exploitation is its ability to retrieve pertinent information on demand. However, the ability to retrieve pertinent information on demand is a difficult problem, since pertinent information must be provided in response to all "reasonable" questions. Furthermore, a "reasonable" question must be defined as any serious question—of obvious or potential significance—posed by persons—with proper training or not—who have scientific, economic, military or other socially important reasons for desiring the answer to the question. In order to provide this type of question-answering potential, a system has been developed by the Center for Documentation and Communication Research at Western Reserve University. What this involves is the encoding of full, informative abstracts in such a way that extremely detailed information is available for search and correlation. Further, the coding system provides an extremely flexible cross-reference system which controls the synonymous or nearly synonymous terminology which may be encountered in the literature. In addition, a special-purpose computer capable of scanning this encoded material has been designed, constructed and tested over a period of three years—leading to the acquisition of a high-speed version of this machine—the General Electric-250 in February, 1961, by Western Reserve University. This equipment will be able to scan 5.4×10^7 symbols per hour and answer ten questions simultaneously.

These methods have been applied to the world's metallurgical literature in a program sponsored by the American Society for Metals. An operational mechanized searching service based on this program has started on January 1, 1960, and is being offered to the membership of ASM at this time. The Center for Documentation and Communication Research is also applying the same methods to other bodies of the literature, as follows:

Electrical Engineering—cooperative project with University of Arizona, sponsored by Fort Huachuca, Ariz.
Diabetes—sponsored by American Diabetes Association
Disease Vector Control—cooperative project with Communicable Disease Center (Atlanta, Ga.), sponsored by National Institutes of Health
Law—cooperative project with Western Reserve University School of Law
Economic and Market Information—program in commercial intelligence sponsored by Rome Air Development Center
Inorganic Chemistry, Solid State Physics, Geology, and Mechanical Engineering—test and evaluation program sponsored by National Science Foundation

We are confident from the results of our test and pilot programs to date that we can meet the criteria for really sophisticated service—at least at the present stage of information requirements.

Administrative Problems

The administrative problems involved in a national center are considerable. The chief difficulties are with regard to the need for servicing from one central point many questions from distant points. In order to provide orderly and rapid service, the various rapid communication means are available. But also, it appears to be necessary to provide a balance between centralized *processing* of information and decentralized *exploitation* where this may be important. Here, it appears that a national center should provide copies of encoded material to various groups throughout the country who may desire either more convenient or more confidential services. For example, the ASM service will provide copies of the world's metallurgical literature, encoded on cards or tape, so that the acquisition of suitable searching equipment will permit various industrial and governmental groups to conduct searches of their own based on the published literature as well as on the encoded confidential reports that they may wish to process themselves.

Other administrative problems involved in operating a national center are the normal accounting problems for accurate and convenient billing for services as well as computation and payment of royalty fees to holders of copyright material used by the Center.

Cooperative Problems

Another problem in operating a center involves that of arranging and policing cooperative arrangements with processors of various segments of the literature, so that excessive duplication of this effort may be avoided. Here, it is important that the "vested interests" in this

field are honored to the extent that they are willing to cooperate in the operation of a national *exploitation* center. There is much work to be done in this regard, since the methods of analysis and encoding of the literature must be digested and learned by the various cooperating groups.

Political Problems

The political problems of operating a national center are somewhat analogous to those of cooperation just discussed. The interest of Senator Humphrey and his Subcommittee on Reorganization (of the Committee on Government Operations) in this field have stemmed from a desire to establish a new cabinet post—involving the formation of a Department of Science. It has been felt that this department should take a strong interest in information services and perhaps even operate a national center in Washington, D. C. The needs of government for information services—although tremendous—would have to be coordinated with those of industry, universities, and other institutions and individuals. The community of interest between governmental and nongovernmental requirements is sufficiently great so that use of the same type of processing of information would be considered advantageous.

However, we find a report issued by the Killian Science Advisory Committee which states:

> Our very progress of science is dependent upon the free flow of scientific information, for the rate of scientific advance is determined in large measure by the speed with which research findings are disseminated among scientists who can use them in further research.
>
> . . . . . .
>
> It is generally agreed . . . that the magnitude and seriousness of the problem are such that a long-term solution requires fundamental research into the problem and widespread application of machine methods and techniques.

In discussing a national center the Killian report continues as follows:

> The proposal to solve existing problems in the field of scientific information by the establishment of a single large operating center, financed wholly or in part by the Federal Government, may have been suggested by the experience of the Soviet Union with its All-Union Institute of Scientific Information. The organization and operation of the Institute implies that the Russians recognize the magnitude and importance of the problem by their decisive and aggressive attempts to meet it. Available evidence indicates that the Institute operates effectively in meeting the needs of Russian science. But, it must not be overlooked that in planning the establishment and operations of the Institute, the Russians could not call upon the services of scientific

> information organizations such as we find already in existence in the private enterprise structure of our country, and which have been in operation many years.
>
> The solution the Russians have developed for meeting their own problems in our judgment *would not* be equally effective in meeting ours. The Russian Institute is organized along the lines that are basically compatible with the organization and administration of research in the Soviet Union, which, of course, is controlled by the Central Government. Our own research efforts are organized and administered very differently, and it is illogical to suppose that a highly centralized organization for the dissemination of research information would serve our purposes equally well. Whatever its faults may be, our present system has developed along the lines of individual initiative and private enterprise that are very basic to our institutions.

The Killian report jumps to conclusions too quickly in the following flat statement:

> From a purely practical point of view, it must be remembered that much of the day-to-day work involved in the dissemination of scientific information—that is, the writing, editing, abstracting, translating, and so on—is done either by scientists or people with technical skills of a very high order. Many of these people perform such chores in addition to their regular scientific work and it is quite conceivable that they could be induced to affiliate themselves on a full-time basis with a centralized agency. Put the matter another way: The case for a Government-operated, highly centralized type of center can be no better defended for scientific information services than it could be for automobile agencies, delicatessens, or barber shops.

However, the Killian report proposes a Science Information Service with the following two broad functions:

> . . . (1) through effective coordination and cooperative effort of public agencies and private organizations to capitalize upon and improve existing facilities and techniques in such a way as to afford immediate relief to short-term problems of a pressing nature; and (2) to encourage and support a fundamental, long-term program of research and development, looking to the application of modern scientific knowledge to the over-all problem through the application of machine techniques and through yet-undiscovered methods.

All this is very well for public consumption, but it was rumored that the Cabinet meetings at which this report was presented led to budget recommendations in this area of only $135,000.

This is a far cry from the public recommendations that were made. However, for various reasons, the amounts actually spent were much greater, due, I presume, to increasing realization during the year of the magnitude of the problem.

Let me hold in opposition to the Killian report the statement made to me in Moscow in March by Dr. A. N. Nesmeyanov, President of the U.S.S.R. Academy of Sciences, when he said that the solution to the scientific information problem is now the most pressing problem whose solution has been undertaken in the U.S.S.R., with a priority far in excess of any other scientific question with which they are coping.

THE FUTURE

The U.S.S.R. has mounted a massive effort to "brainpick" the world's recorded literature in order to assure a more effective scientific and technical effort on their part.

We, too, *must* go in this direction. Our "brainpicking," or exploitation efforts, have been desultory. The various groups working on this problem are working largely independently with little coordination or cooperation. It appears that despite the national importance of the solution to this problem the final determination will be made—as always in the United States—in the market-place! It is an efficient procedure—the way of the rich—but regardless, we must now get on with it!

BIBLIOGRAPHY

ALLEN KENT, "A Machine that Does Research," *Harper's Magazine*, April 1959, pp. 67 ff.

ALLEN KENT AND A. S. IBERALL, "Soviet Documentation—A Trip Report," *American Documentation*, Vol. 10, pp. 1–19 (January 1959).

J. W. PERRY AND ALLEN KENT, "Tools for Machine Literature Searching," New York, Interscience Publishers, 1958.

ALLEN KENT AND J. W. PERRY, "The Storage and Retrieval of Non-numerical Data in Large and Complex Documentation Systems," Paper No. 59–IRD–5, ASME Publication, New York, 1959.

ALLEN KENT AND J. W. PERRY, "The Information Center of Tomorrow," Paper No. 58–MD–13, ASME Publication, New York, 1958.

ALLEN KENT AND J. W. PERRY, "Centralized Information Service," New York, Interscience Publishers, 1958.

ALLEN KENT AND OTHERS, "Machine Searching of Metallurgical Literature," *Metal Progress*, Vol. 71, No. 2, pp. 71–75 (February 1957).

ALLEN KENT, "Machine Literature Searching and Translation—An Analytical Review," summary paper presented before International Conference for Standards on a Common Language for Machine Searching and Translation, Cleveland, Ohio, September 6–12, 1959. (Chapter 1 of "Information Retrieval and Machine Translation," edited by A. Kent, Volume I, New York, Interscience Publishers, in press.)

THE FRANKLIN INSTITUTE

COMMITTEE ON SCIENCE AND THE ARTS

(Abstract of Proceedings of Stated Meeting held Wednesday, April 13, 1960.)

HALL OF THE COMMITTEE,
PHILADELPHIA, APRIL 13, 1960.

MR. J. PHILIP EVANS in the Chair.

The following reports were presented for final action:

No. 3324: The Airborne Magnetometer.

This report recommended the award of a John Price Wetherill Medal to Victor V. Vacquier, of La Jolla, California, "In recognition of his important invention of the sensitive saturating core magnetometer element, and his development of the first practical airborne magnetometer, providing the fundamental principles which are basic to the later developments.".

No. 3367: Work of R. R. Moore in Fatigue Testing of Materials.

This report recommended the award of a Certificate of Merit to Richard R. Moore, of Drexel Hill, Pennsylvania, "In consideration of his pioneering achievements in the field of fatigue testing of materials."

No. 3373: The Franklin Medal.

This report recommended the award of The Franklin Medal to Roger Adams, of Urbana, Illinois, "In recognition of his many contributions to the science of organic chemistry, particularly his widely useful method of catalytic hydrogenation, his comprehensive investigations of alkaloids, and his systematic studies of molecular structure; and because of his eminent achievements as a teacher, as an author and editor, and as a leader in the professional organizations of science."

No. 3374: Brown Medal.

This report recommended the award of the Frank P. Brown Medal to Richard Buckminster Fuller, of Forrest Hills, New York, "In consideration of his inventive conception of the Geodesic Domes, the development of the associated mathematical derivations and the resolution of the theoretical principles to practice as a vastly more efficient and economic means of providing full-span shelter cover against the elements, without need for interior supports."

D. S. FAHRNEY,
Secretary to Committee

(Abstract of Proceedings of Stated Meeting held Wednesday, May 11, 1960.)

HALL OF THE COMMITTEE,
PHILADELPHIA, MAY 11, 1960.

MR. J. PHILIP EVANS in the Chair.

The following reports were presented for final action:

No. 3337: Electron Probe Microanalyser.

This report recommended the award of a John Price Wetherill Medal to Raymond Barnard Castaing, of Paris, France, "In consideration of his development of the electron probe microanalyser, which has made possible the metallurgical analysis of a sample, not only as to structure, but also quantitatively the chemical analysis of its structural components down to particles of one micron in diameter."

No. 3369: Work of W. Edward Chamberlain.

This report recommended the award of an Edward Longstreth Medal to W. Edward Chamberlain, of Washington, D. C., "For his continuous and effective analyses and applications of physics and engineering to the techniques of radiology; for his keen insight into all of the diverse phenomena contributing to the fluoroscopic process; and for his stimulation of the development of new devices."

No. 3375: Ballantine Medal.

This report recommended the award of the Stuart Ballantine Medal to Harry Nyquist, of Milburn, New Jersey, "For his theoretical analyses and practical inventions in the field of communications during the past forty years including, particularly, his original work in the theories of telegraph transmission, thermal noise in electric conductors, and in the theory of feedback systems."

D. S. FAHRNEY,
Secretary to Committee

(Abstract of Proceedings of Stated Meeting held Wednesday, June 8, 1960.)

HALL OF THE COMMITTEE,
PHILADELPHIA, JUNE 8, 1960.

MR. J. PHILIP EVANS in the Chair.

The following reports were presented for final action:

No. 3317: Work of C. S. Draper on Inertial Navigation.

This report recommended the award of a Howard N. Potts Medal to Charles Stark Draper, of Cambridge, Massachusetts, "In consideration of his substantial and significant contributions to the science of inertial navigation, which have resulted in outstanding advances in the accuracy of navigation below and on the surface of the ocean, through the air and into space."

No. 3356: Self-Priming Centrifugal Pump.

This report recommended the award of the Certificate of Merit to Harry E. La Bour, of Elkhart, Indiana, "In consideration of his conception and development of the Self-Priming Centrifugal Pump which has been an outstanding contribution to the pump industry."

No. 3366: Work of A. Nadai.

This report recommended the award of an Elliott Cresson Medal to Arpad Ludwig Nadai, of Pittsburgh, Pennsylvania, "For his pioneering work in the field of Elasticity of materials and

in Plastic Flow through many research contributions, for important educational activities, and for his authoritative treatises."

No. 3376: Ballantine Medal.

This report recommended the award of a Stuart Ballantine Medal each to Rudolf Kompfner, of Far Hills, New Jersey, and John R. Pierce, of Berkeley Heights, New Jersey, "For the conception and invention of the traveling wave tube amplifier, a device based on a new principle of continuous interaction between an electromagnetic wave and an electron beam, that removes the frequency bandwidth limitation of electronic tubes; for the formulation of the fundamental theory of its operation and for its development as an important instrumentality in commercial and military communications."

No. 3377: Henderson Medal.

This report recommended the award of a George R. Henderson Medal each to Herbert Sammons and Ernest Chatterton, of London, England, "In consideration of their outstanding achievements, the invention, design and development of the 'Deltic' Diesel-electric drive for railway locomotives."

No. 3378: Clark Medal.

This report recommended the award of a Walton Clark Medal each to Robert W. Cook and Lee F. McBride, of Decatur, Indiana, "For their contribution to public welfare and safety by the development of a successful leak-sealing system for gas mains."

No. 3379: Levy Medal.

This report recommended the award of a Louis E. Levy Medal each to Duane T. McRuer, of Inglewood, California, and Ezra S. Krendel, of Swarthmore, Pennsylvania, "In recognition of their outstanding paper, 'The Human Operator as a Servo System Element,' appearing in the May and June, 1959, issues of the JOURNAL OF THE FRANKLIN INSTITUTE."

No. 3347: Work of W. F. G. Swann.

This report recommended the award of an Elliott Cresson Medal to William Francis Gray Swann, of Swarthmore, Pennsylvania, "For his many investigations in numerous areas of physical science and unrelated subjects that have broadened and deepened an understanding and a meaning of the universe and life therein, and more particularly for his profound and significant studies and researches in the field of cosmic radiation."

D. S. FAHRNEY,
Secretary to Committee

READER OPINION *

Cosmic Thunderstorms.—In a recent article in this JOURNAL,[1] and in other contributions,[2] the writer showed that his electrical discharge theory of those magnetic storms, which follow outbursts at the sun's surface by periods of the order of a day, led to the surprising conclusion that temperatures of the order of 400,000,000°K must be reached in these discharges somewhere between the sun and the earth.

In another presentation[3] of this work the writer emphasized, however, that this theoretical conclusion need not surprise us too much, since temperatures of over one million degrees are observed in the solar corona. Temperatures of this order too have been obtained in electrical discharges in the laboratory. Furthermore, the point was made that other conditions exist in which a considerable increase in temperature occurs when these discharges are propagated down a density gradient in stellar atmospheres. In the combination-spectra stars, for example, in which the density gradient is very much less than it is in the solar atmosphere, the discharge temperature similarly increases from about 5000°K or 10,000°K, at the start of the outburst, to a final value of over a million degrees absolute. In the article itself[1] it was pointed out that these long atmospheric electrical discharges act as "energy pumps," liberating in regions of low gas density the electric field energy which had been generated lower down in the stars' atmospheres.

Nevertheless this deduction from the theory, while put forward with some confidence in view of the theory's mounting success in accounting for other gas movements in stellar and galactic atmospheres, was also recognized as a possible source of criticism.

It was therefore reassuring when the writer's attention was drawn to a reference[4] to the work of the U. S. Naval Scientists. In their study of the phenomena accompanying solar flares, which, it was suggested earlier[5] are electrical discharges, these Scientists have found evidence outside the earth's atmosphere of X-rays having energies of 80,000 eV, and therefore evidence of the existence of temperatures of 100,000,000°K, in good agreement with the theoretical predictions.

C. E. R. BRUCE
The Electrical Research Association Laboratory
Leatherhead, England

* EDITOR'S NOTE: For this new section of the JOURNAL, readers are invited to submit comments on papers published in previous issues. Contributions should indicate clearly (by volume, page number, title and author) what paper is under discussion; they should be specific, not general, in nature; and they should be limited to approximately 500 words.

[1] C. E. R. BRUCE, "Cosmic Thunderstorms," this JOURNAL, Vol. 268, p. 425 (1959).

[2] C. E. R. BRUCE, *Nature*, Vol. 184, p. 2004 (1959).

[3] C. E. R. BRUCE, *Comptes Rendus*, Vol. 250, p. 61 (1960).

[4] Armour Research Foundation, "Industrial Research News Letter," November, 1959, Item 18.

[5] C. E. R. BRUCE, *Engineer*, Vol. 208, p. 864 (1959).

BOOK REVIEWS

CHEMISTRY FOR SANITARY ENGINEERS, by Clair N. Sawyer. 367 pages, diagrams, 6 × 9 in. New York, McGraw-Hill Book Co., Inc., 1960. Price, $9.50.

Books on the subjects not already overwritten are always welcomed for filling the gap. *Chemistry for Sanitary Engineers* fills a vital need but its acceptance will not be based on this alone. Dr. Sawyer, one of sanitary engineering's foremost scientists, has prepared a readable, concise text for the student and a handy reference for the professional engineer and plant operator.

Except for advanced degree candidates, there is little opportunity for undergraduates in the field to obtain training in all the various chemical subjects so integral to sanitary engineering. This book serves to effect such coverage. It is divided into two parts: the first nine chapters summarize the basic principles of such topics as biochemistry, radiochemistry and colloidal chemistry; the next 24 chapters are devoted to various sanitary chemical tests for water, industrial waste water, and sewage. The experienced sanitary chemist will find these sections to be complimentary to the *ASTM Manual on Industrial Water* and the *Standard Methods*. Actual analytical procedures are not presented but the importance of the test, the physical-chemical or biological principles behind the methods, and the applications are described. The clarity of the presentation on pH, acidity, and alkalinity merits review of these fundamental but frequently misunderstood relationships by the experienced engineer as well as the student. The description of the role of nitrogen is another example of well developed but easy-to-follow presentation.

There are relatively few references. Future editions might be enhanced by addition of up-to-date bibliographies for use by the advanced student or practicing engineer. Instrumental chemical analytical procedures such as those employing visual and infrared spectrographic and chromatographic methods might also be included. Their applications, especially in research work, are becoming more important in the field.

In this reviewer's opinion, *Chemistry for Sanitary Engineers* will be a worthwhile addition to the library of those in the fields of water or waste treatment, water-side corrosion or sewage treatment.

ROBERT A. BAKER
The Franklin Institute Laboratories

DIGITAL AND SAMPLED-DATA CONTROL SYSTEMS, by Julius T. Tou. 631 pages, diagrams, 6 × 9 in. New York, McGraw-Hill Book Co., Inc., 1959. Price, $15.00.

This excellent book is an important accession to the few books on sampled-data control systems recently published. It will be useful both as a text book and as a reference.

Most of the chapters begin with an introduction, giving the reader some idea of what to expect. A conclusion at the end of many chapters emphasizes the main points preented. The chapter arrangement is extremely logical, and the material within the chapters shows much more than just casual organization. The writing is clear and concise and sufficient tables and diagrams are included to enhance presentation greatly.

The first two chapters are of introductory nature. The use of root locus for design of continuous systems has been skillfully expounded.

The next four chapters deal with system analysis. The theory and application of both Z and modified Z transform methods are given thorough treatment. System design, error analysis, statistical method, and finite sampling are covered in the remaining three chapters.

Multirate sampled systems are discussed. One chapter is devoted to the description of the analog-digital conversion principle. Nyquist and Bode plots as well as root locus method in sampled-data systems have also been appropriately presented.

Many problems given in the appendix are extremely helpful to all students. A ready

source of references is available through the appendix at the end of the book.

On the whole, the author has done an excellent job of presenting digital and sampled-data control systems in a single volume and summarizing the great amount of literature published in this field.

Y. H. KU AND D. W. C. SHEN
The Moore School of Electrical Engineering
University of Pennsylvania

INTRODUCTION TO THE MECHANICS OF THE SOLAR SYSTEM, by Rudolf Kurth. 177 pages, 5½ × 8½ in. New York, Pergamon Press, 1959. Price, $6.50.

The aims for this book as stated in the Preface are laudable, but sometimes the pursuit of those aims descends to triviality. For instance (page v) "Nearness to natural reality —as is seen, for example, in the occasional use of such names as Jupiter and Saturn." When attainable, "Insight through simplicity" is also a worthwhile aim; in particular, improvement of results by extension and refinement of "*the same* [italics are both Dr. Kurth's and mine] method by which they were derived" is excellent in the abstract, but obviously in the nature of things not infrequently impracticable.

The organization of the book into four parts or chapters—Kinematics of a Single Planet, Dynamics of a Single Planet, Dynamics of the Planetary System, The Planets and The Moon as Rigid Bodies—is well designed for the overview Dr. Kurth is trying to give. The exercises given at the ends of the chapters show the same curious mixture of high ideal and sudden descent to triviality that this reviewer felt sometimes in the text; what results is not simply variety of problem, but inconsistency of academic level at least as far as the American student is concerned. The apparent distaste for "getting his hands dirty" with computing which Dr. Kurth displays in the Preface shows up rather strongly in his choice of exercises, but this is at least consistency, and the bulk of the exercises seem well selected.

This reviewer did not, of course, undertake to check every step in the text, but was impressed with the unusually small number of errors of detail noted in one fast and one more careful reading; the minor one at the bottom of page 76 is actually the lone case uncovered. There occur, however, here and there, cases where opinion is cloaked in the phraseology of fact; the statement on page 43, that "Had he done so, physical science and philosophy would have developed in a completely different way" seems to belong more properly in a discussion in a seminar room than in the cold print of a textbook. In fact, the contrast between this and the honest simplicity of (page 147) "I have no practical experience in the integration of the perturbation equations. My proposals for the computation of the solution, therefore, indicate only what I suppose should be tried" seems to provide a rather good measure of the book.

In this reviewer s opinion, any one of three or four already established texts will provide a more balanced introduction of the American undergraduate and early graduate student to celestial mechanics, but *after* an initiation through a rather more practical approach, that same student would gain some breadth of horizon by reading Dr. Kurth's volume.

JOHN E. MERRILL
The Franklin Institute Laboratories

INTRODUCTION TO MATRIX ANALYSIS, by Richard Bellman. 328 pages, 6 × 9 in. New York, McGraw-Hill Book Co., Inc., 1960. Price, $10.00.

The importance of matrix methods in pure mathematics, mathematical physics, electrical engineering, statistics, etc. has been increasing steadily in the last two decades and the literature on this subject has been increasing at an even faster rate. R. Bellman's book, the first of a McGraw-Hill series on matrices, claims to be an introduction to the subject, but is in fact much more than that, as the contents immediately indicate. The first eight chapters (140 pages) dealing with symmetric matrices constitute the introduction, immediately followed by a chapter on Dynamic Programming. Chapters 10 to 13 are devoted to the application of matrices to the solution of differential equations, including stability theory. In the last three chapters, the theory of position matrices is treated with its application to probability theory and mathematical economics.

The book is written with great pedagogic skill, with numerous short exercises illustrat-

ing the abstract treatment. Many good problems at the end of each chapter extend the treatment in the text. A carefully selected bibliography follows each chapter, and the contents of many references are briefly discussed.

This reviewer feels that the application of matrix theory to quantum mechanics has been relatively neglected. Otherwise this is an excellent introduction to the subject of matrices, and if the subsequent volumes meet the standards of this first volume, no doubt the series will be a very useful and successful one.

HENRI AMAR
Temple University and
The Franklin Institute Laboratories

REGRESSION ANALYSIS, by E. J. Williams. 214 pages, diagrams, 6 × 9 in. New York, John Wiley & Sons, Inc., 1959. Price, $7.50.

This publication marks the appearance on the scene of yet another book on curve fitting. This particular subject appears to be capturing the fancy of those who write texts on statistics for the non-statistician. To keep the reader from being misled, the phrase "for the non-statistician" should not be interpreted as "statistics made easy."

The text, here under discussion, is quite complete covering linear and curvilinear regression, covariance analysis ("curve fitting" with qualitative data) and discriminant analysis (linear functions which best discriminate between similar systems). The latter two are usually not covered in these specialized texts.

The book is well written and organized, and should prove a valuable reference to those with a firm background in statistical theory, or as the author claims—"the reader is familiar with or has ready access to these methods." Because this type of background is valuable, if not necessary, it is doubtful whether the text, because of its special nature, can serve "research workers in the experimental field" as readily as the author hopes. This statement can be generally applied to almost any of the statistical texts in existence.

ALOYSIUS J. POLANECZKY
The Franklin Institute Laboratories

ADVANCES IN SPACE SCIENCE, VOLUME I, edited by Frederick I. Ordway, III. 412 pages, diagrams, 6 × 9 in. New York, Academic Press, 1959. Price, $12.00.

Six monographs have been incorporated into *Advances in Space Sciences*. The publishers were wise in inviting Mr. Ordway to edit the series, for his wide knowledge of space travel assured the selection of the most topical and authoritative pieces for this book. Because of the tremendous interest and work being done in bioastronautics, three of the six pieces are concerned with this problem.

The opening section by Derek F. Lawden, on "Interplanetary Rocket Trajectories," is an impressive, mathematical treatment of the problem of getting into orbit and in motion between the planets. This is a highly technical treatment suitable for the engineer and worker in the field.

Drs. J. R. Pierce and C. C. Cutler treat communications in space. In a definitive study, they explore various types of satellites to carry out certain communications tasks they set up. They compare the many forms of passive and active satellite relay configurations and assess each for potential, simplicity, weight and in some cases cost. They discuss the feasibility of interstellar and interplanetary communications and conclude the first is virtually impossible and the second is quite possible. (Within the next few months we will be receiving signals from Pioneer V at a distance of from 50 to 75 million miles.)

John H. Huth, of the Rand Corporation, in his section, analyzes power sources for space vehicles. In this section we find a complete run-down on the various systems—both in use and contemplated—to provide this power. He derives the power requirements for the present and the immediate future and then launches into the conversion techniques to yield the necessary power. He explores the electrochemical approach, the use of turbo generators, magnetohydrodynamic conversion, the use of thermocouple, thermionic devices, the silicon solar cell and, finally, fission-electric conversion.

"Manned Space Cabin Systems," by Eugene B. Konecci of Douglas, is the next chapter in this book. This 100-page section represents the most complete treatment this reviewer has ever seen concerning the prob-

lems of man in a space cabin. There are three subsections dealing with the human body. In this are delineated the man-machine relationships and the limitations of man as an organism. The major subsection deals with the life-support system, that is, the cabin with its atmosphere, temperature, susceptibility to puncture by meteorites, and potential to fulfill human requirements. Mr. Konecei has supplemented this with an extensive reference list.

Dr. Hermann J. Schaefer uses the next section to treat "Radiation and Man in Space." Again we are exposed to a definitive treatment of radiations in space and their effect on man. The author first discusses primary cosmic rays, describing the biological effect on man of various types of nuclei. He speaks of the increase in cosmic rays with solar flares and then details information on the Van Allen belts including the correlation between intensity and radiation dosage.

The last section by Prof. Robert G. Tischer is on "Nutrition in Space Flight." After an introduction on work being done on the problem of space feeding, the author gets down to specifics. What are the nutritional requirements of a man and how can *Alga Chlorella Pyrenoedosa* become part of a closed cycle system? It appears inevitable that eventually man will be in space for extended periods and then algae must become part of the picture.

The book concludes with a "Decimal Classification System for Astronautics," by H. H. Koelle.

Advances in Space Science is a well written, definitive book on several of the most pressing problems in space travel.

I. M. LEVITT
The Fels Planetarium

SOLID STATE PHYSICS, VOLUME 9, edited by Frederick Seitz and David Turnbull. 548 pages, diagrams, 6 × 9 in. New York, Academic Press, 1959. Price, $14.50.

The steady flow of surveys and reviews of various topics in solid state physics continues, as this area of research grows at a remarkable rate. This, the ninth volume of the series edited by Seitz and Turnbull, meets the generally high standards of the preceding numbers. In view of the diversity of the topics covered, this reviewer feels that a brief and comprehensive description of the contents would be more useful to the reader than some general comments. The 525 pages are fairly evenly divided among the following seven topics.

1. *The Electronic Spectra of Aromatic Molecular Crystals* (H. C. Wolf) surveys past and present investigations on the spectra of aromatic crystals. The theory of these spectra (in particular a quantitative treatment of displacement and splitting) is then compared with the experimental data.

2. *Polar Semiconductors* (W. W. Scanlon). This exceptionally lucid and methodical presentation begins with the analysis of crystalline and physicochemical properties of polar crystals and ends with a thorough examination of the electrical and optical properties of the alloys PbS, PbSe, PbTe.

3. *Static Electrification of Solids* (D. J. Montgomery). The author points out that the understanding of this old and very important subject is still insufficient especially when compared to that of other areas of Solid State Physics. He confines himself mostly to the transfer of change between Solids. A brief section on the directions of future work will be found valuable.

4. *The Interdependence of Solid State Physics and Angular Distribution of Nuclear Radiations* (Ernst Heer and Th. D. Novey). The unperturbed angular correlation of Nuclear Radiation is first exposed and followed by the study of extranuclear effects involving paramagnetic and non-paramagnetic atoms. Experimental techniques are considered. A brief section is devoted to this field as a tool in Solid State Physics.

5. *Oscillatory Behavior of Magnetic Susceptibility and Electronic Conductivity* (A. H. Kahn and H. P. R. Frederikse). This is an excellent review of the DeHaas-Van Alphen Effect with exceptional emphasis in Bismuth. Other metals and semiconductors are briefly considered.

6. *Heterogeneities in Solid Solutions* (Andre Guinier). One of the leading experts discusses homogeneous, disordered solutions, in equilibrium and otherwise, the influence of lattice defects on the kinetics of zone development.

7. *Electronic Spectra of Molecules and Ions in Crystals* Part II (D. C. McCluse). This second part devoted to the spectra of ions exhibits the same qualities of thoroughness and clarity as the first part on Molecules.

HENRI AMAR
The Franklin Institute Laboratories and Temple University

V.2
196

OPEN CHANNEL HYDRAULICS, by Ven Te Chow. 680 pages, diagrams, 6 × 9 in. New York, McGraw-Hill Book Co., Inc., 1959. Price, $17.00.

Periodic summaries are a necessary part of the documentation cycle of any growing field of research, if the training and indoctrination of students and workers are to be facilitated. Especially when the field involved is a big one, these summaries tend to be group efforts, usually *ad hoc* collections of specialist contributions which almost necessarily show deficiencies of organization and an uneven level of presentation. Although single-author efforts do sometimes amount only to a string of abstracts with enough text to provide continuity, the opportunity to apply unity of treatment and purpose usually makes for a better book, and can result in the finest kind of technical literature. Ven Te Chow's *Open Channel Hydraulics* is a beautiful example of just how good a one-man summary can be.

This book is remarkable for its comprehensive scope, logical organization, lucid writing, and steadfast adherence to a single pedagogical objective. The labor that went into it must have been truly heroic.

The author has covered virtually all practical cases of open channel flow, in the process drawing upon the literature of several languages, including the Russion for his sources. Over 800 specific references are cited, most of them published in the last ten or fifteen years.

Five major sections, divided into twenty chapters totaling over 600 pages, cover basic principles, uniform flow, gradually varied flow, rapidly varied flow, and unsteady flow. The author has aimed at a balance between theory and practice and has successfully held to it throughout. Mathematical elegance is de-emphasized in favor of producing a readable, teachable, usable book which would have as wide an audience as possible among both students and engineers. The examples at the end of each chapter, for instance, are about half of the concept reinforcement type, involving demonstrations or proofs, and about half design problems serving to simulate practice. In the text, theory is simplified but adequate. Energy and momentum principles are used exclusively for purposes of conceptual economy and unity. Heavy emphasis is placed on practical methods of calculation in the treatment of design problems.

The book is full of charts, diagrams, curves, tables and nomograms of excellent quality. Doubtless the cost of preparing these aids accounts for a good part of the exceptionally high price of the book. It is unfortunate, in view of the overall excellence of the printing, that some of the halftones seem to be so poor.

Professor Chow's book is the first summary in the field to appear in English in nearly twenty years. It is safe to predict that another decade will go by at least before another treatise of equal quality appears.

WILLIAM C. YAGER
The Franklin Institute Laboratories

STEROIDS, by L. F. Fieser and M. Fieser. 945 pages, diagrams, 6 × 9 in. New York, Reinhold Publishing Corp., 1959. Price, $18.00.

As an organic chemist not working in the steroid field, the reviewer approached this book cautiously at first, but with the assurance that the energetic and clear style of the Fiesers would carry him through an unfamiliar field. But caution, in this case, proved far from necessary. The authors have succeeded in the difficult task of writing a book which is stimulating to the non-steroid organic chemist, but one which appears complete and up-to-date for the specialist as well.

Several points about the book are particularly noteworthy: the inclusion of literature through early 1959; a concise effective orienting survey (Chapter 1); historical information which gives the reader insight into the development of the field. Chapters 5–10 discuss physical methods of characterization, oxidation, enes and ols, ketones, displacements and rearrangements, stereochemical correlations and conventions, in a manner applicable to all of organic chemistry using steroids as examples. The other chapters deal primarily with the bile acids, sterols, Vitamin D, the hormones, sapogenins and alkaloids. The rapid growth in the field has made it extremely difficult for anyone but the specialist to attempt to keep abreast of the developments and the significance and relevance to other chemical problems. The book performs this very valuable function and performs it well.

As to its value to the specialist, colleagues working in the field assure me that the book is both exciting and indispensable to them.

Having been able to satisfy both steroid and non-steroid workers, the Fiesers and the publishers are to be thanked for a most valuable contribution to the literature.

MORTIMER M. LABES
The Franklin Institute Laboratories

ANALYTICAL AND CANONICAL FORMALISM IN PHYSICS, by Andre Mercier. 222 pages, diagrams, 6 × 8¾ in. New York and London, Interscience Publishers, Inc., 1959. Price, $6.75.

The second quantization in quantum field theory and quantum electrodynamics requires as a preliminary step the reformulation of the problem, using Hamiltonian and Lagrangian densities, and canonical variables. This book will do much to clarify for the student the relationships between the various canonical formalisms, and to remind him that they are strictly classical formalisms, although they find their most important applications in quantum field theory.

In the first two chapters the author introduces the Lagrangian formalism and Lagrangian densities to be used in the treatment of the canonical formalism (third chapter). The methods used are illustrated in the fourth chapter where a canonical form of electrodynamics is derived. The rest of the book (3 chapters) deals with Hamiltonian densities, canonical transformation in general, and those with vanishing Jacobian style. Each chapter is preceded by a detailed summary of contents. Since the book is mostly didactic, more examples solved in the text would have been a valuable addition. As it is, this book is the first of its kind to give a unified and coherent exposé of transformations and of the various canonical formalisms.

HENRI AMAR
Temple University and
The Franklin Institute Laboratories

BOOK NOTES

PLASMA DYNAMICS, by F. H. Clauser. 369 pages, diagrams, 6×9 in. Reading (Mass.), Addison-Wesley Publishing Co., Inc., 1960. Price, $12.50.

Plasma dynamics (also called magneto hydrodynamics, hydromagnetics, Hg-dynamics and magnetoaerodynamics) deals with the interaction of deformable conducting materials and the electromagnetic field. It is, therefore, of vital importance to researchers in the fields of astrophysics, fluid mechanics, thermonuclear physics, statistical mechanics and aerodynamics. The present volume is based on an international symposium held in June 1958, and, in the light of the rapidly changing picture in this field, perhaps the most serious criticism of the book is the 18-month delay between the symposium and publication of the edited proceedings. This delay is understandable (a large portion of the text is composed of condensed versions of oral discussions) but regrettable. The nine chapters cover experimental research, thermonuclear fusion, gaseous electronics phenomena, dynamics of electron beams, statistical plasma mechanics, continuum plasma dynamics, flight magnetohydrodynamics, solar and interplanetary magnetohydrodynamics, and cosmical magnetohydrodynamics. In spite of the publication delay, this is the most inclusive and authoritative book yet to appear on this important subject.

QUANTITATIVE MOLECULAR SPECTROSCOPY AND GAS EMISSIVITIES, by S. S. Penner. 587 pages, diagrams, 6×9 in. Reading (Mass.), Addison-Wesley Publishing Co., Inc., 1959. Price, $15.00.

The author has prepared this text for "students in applied science as an introduction to the literature on a class of radiation problems which fall . . . into the temperature range intermediate between that of interest to the molecular spectroscopist . . . and the astrophysicist." Both theory and experimental studies are included and emphasis is on those problems which are encountered in the development of modern propulsion devices. An interesting feature is the omission of material which is found in Herzberg's compilations—it is refreshing to find an author who does not think it necessary to summarize previous work in his own text. This new text should find wide acceptance, for it is well written, practical and timely.

THE PROTON IN CHEMISTRY, by R. P. Bell. 223 pages, diagrams, 6 × 9¼ in. Ithaca, Cornell University Press, 1959. Price, $4.75.

The author bases this book on the series of lectures given at Cornell University during 1958, when he was George Fisher Baker Non-resident Lecturer. Ten in number, the lectures range from "The Qualitative Nature of Acids and Bases" to "Isotope Effects in Acid-Base Reactions." In "Concentrated Solutions of Acids and Bases," the author describes recent work in this field; discussion of the reactions kinetics of these concentrated solutions has been omitted. Also, two chapters are devoted to the relations between molecular structure and acid-base properties.

CATALYSIS, VOLUME 7, edited by Paul H. Emmett. 378 pages, diagrams, 6 × 9 in. New York, Reinhold Publishing Corp., 1960. Price, $13.50.

Completing the series on Catalysis, Volume 7 provides research chemists with a soundly conceived reference source on the physical chemistry of catalysis and catalytic processing. Four chapters are included on catalytic oxidation, one on hydration-dehydration and one on cracking catalysts. This final volume, together with the other volumes of the series, should prove of tremendous value to physical chemists and to every industry in which catalysis is involved.

FLUID MECHANICS, by L. D. Landau and E. M. Lifshitz. 536 pages, diagrams, 6 × 9 in. Reading (Mass.), Addison-Wesley Publishing Co., Inc., 1960. Price, $14.50. (Available in U. S. and Canada only.)

This is Volume 6 in the Course of Theoretical Physics and is translated from the Russian by J. B. Sykes and W. H. Reid. This text describes fluid mechanics as a branch of theoretical physics, and the authors have tried to develop as fully as possible all matters of physical interest to give the clearest possible picture of the phenomena and their interrelation. They discuss neither approximate methods of calculation in fluid mechanics, nor empirical theories devoid of physical significance. On the other hand, some topics not usually found in textbooks on the subject are treated, such as the theory of heat transfer and diffusion in fluids; acoustics; the theory of combustion; the dynamics of superfluids; and relativistic fluid dynamics. While it has not been possible for the authors to incorporate in this English-language edition all new results in the field which have taken place since the Russian edition was published, they have added to the contents of the original edition one new chapter, that on the general theory of hydrodynamic fluctuations.

MATHEMATICAL METHODS FOR DIGITAL COMPUTERS, edited by Anthony Ralston and Herbert S. Wilf. 293 pages, diagrams, 7 × 10¼ in. New York, John Wiley & Sons, Inc., 1960. Price, $9.00.

Each of the twenty-six chapters (all by different contributers) in this reference work describes the process which takes a problem from its formulation up to, but not including, the coding for a digital computer. In this way, the book serves to show the interrelation between analysis of the problem, computer capabilities and limitations, and coding procedures. Each chapter gives the purpose of the program, a mathematical analysis of the specific problem, the calculation procedures to be used, a detailed flow chart, a description of the flow chart, the memory requirements, estimated running time, and a list of references. Examples of the methods covered are the Gauss-Seidel Method, Conjugate Gradient Method, Jacobi Method, Monte Carlo methods, etc.

PRINCIPLES OF CHEMICAL THERMODYNAMICS, by Charles E. Reid. 306 pages, diagrams, 6 × 9 in. New York, Reinhold Publishing Corp., 1960. Price: $6.00 (college); $7.80 (trade).

The author has prepared this new text on chemical thermodynamics because he felt that none of the numerous existing texts adequately presented the subject. In his version: entropy is treated with no attempt to make it dependent on heat engines; a chapter on statistical thermodynamics is included; and he has used terminology and symbolism which differ somewhat from customary usage. The text is intended as an introductory thermodynamics course for graduate chemistry students.

POLYESTER RESINS, by John R. Lawrence. 251 pages, illustrations, 5 × 7½ in. New York, Reinhold Publishing Corp., 1960. Price, $5.75.

This book gives a comprehensive account of the relatively new polyester resin industry, covering all aspects from historic background, chemistry, curing mechanisms and manufacturing, to techniques used in processing the resins for use in reinforced plastics, casting and coating applications. Uses for the versatile polyesters fall in a broad spectrum from boats to missile bodies, and the author has presented them in a concise form. Special emphasis is given to formulating resins for special properties such as weathering and flame resistance. The appendix includes a tabulation of the basic characteristics of standard resins manufactured by major resin suppliers. Numerous tabulations, figures and illustrations contribute to the book's usefulness to technical and non-technical readers in fields both directly and indirectly associated with the use of polyester resins.

COMMUNICATION IN BUSINESS AND INDUSTRY, by William M. Schutte and Erwin R. Steinberg. 393 pages, 6½ × 9¼ in. New York, Henry Holt and Company, 1960. Price, $4.75 (text).

This practical book guides the businessman and the student in the development of basic communication principles. Major attention is given to the problems in communication such as reading and listening, handling ideas, organizing and proper use of punctuation. The author's thorough coverage of the subject establishes a foundation for the eventual mastery of communication skills applicable to a wide variety of business situations. The concluding chapters demonstrate how these skills can be exercised in such practical matters as writing the long letter or report, speaking in public, writing persuasively and leading a conference group. In addition, some chapters contain materials for rewriting which are taken directly from actual business communications.

THE MAGNETODYNAMICS OF CONDUCTING FLUIDS, edited by Daniel Bershader. 145 pages, diagrams, 6 × 9 in. Stanford, Stanford University Press, 1959. Price, $4.50.

In this book are published seven major papers presented at a symposium held at the Palo Alto research laboratory of the Lockheed Missiles and Space Division in November, 1958. The subject matter is concerned with the macro-dynamics of conducting fluids. Two papers deal with the propagation of waves; one paper explores the electromagnetic interaction with aerodynamic flows; one deals with flow and stability problems; one describes investigations of supersonic flow; one deals with the behavior of shock tubes; and one covers flow past a flat plate. The papers give evidence of the real progress made in recent years.

PROGRESS IN SEMICONDUCTORS, VOLUME 4, edited by A. F. Gibson, F. A. Kröger and R. E. Burgess. 291 pages, diagrams and plates 6 × 9 in. New York, John Wiley & Sons, Inc., 1960. Price, $10.50.

Volume 4 of this annual series is the first to include contributions from Japan, Germany, Holland and Russia. In keeping with the purpose of the series, Vol. 4 contains both general and specialized papers. The eight papers in the current volume are: Negative Effective Masses in Semiconductors by H. Krömer (Germany); Oxidation Phenomena on Germanium Surfaces by M. Green (U. S. A.); Theory of Avalanche Multiplication in Nonpolar Semiconductors by J. Yamashita (Japan); Internal Field Emission by A. G. Chynoweth (U. S. A.); Noise in Semiconductors by D. Sautter (Germany); The Electrical Effects of Dislocations in Semiconductors by W. Bardsley (England); Dielectric Properties of Solids in Relation to Imperfections by J. Volger (Holland); and Non-crystalline, Amorphous, and Liquid Electronic Semiconductors by A. F. Joffe and A. R. Regal (Russia).

MAGNESIUM AND ITS ALLOYS, by C. Sheldon Roberts. 230 pages, illustrations, 6 × 9 in. New York, John Wiley & Sons, Inc., 1960. Price, $9.00.

The author, head of the Materials Section of the Fairchild Semiconductor Corp., is well

qualified to prepare a reference work on magnesium and its alloys. He emphasizes the present state of the art, with little historical detail; he also avoids duplication of material dealing with design, fabrication and finishing techniques. The interested reader may look these up, using the extensive lists of references included in the book. The nine chapters, about equally divided between theory and practice, cover physical and chemical properties, alloy theory, deformation, extraction and refining processes, wrought and casting alloy technology, and such phenomena as diffusion, recrystallization and grain growth, and precipitation. The chapter on Magnesium Alloy Systems contains concise data, in graph form, on thirty-three binary and ternary magnesium alloy systems.

RADIATION PYROMETRY AND ITS UNDERLYING PRINCIPLES OF RADIANT HEAT TRANSFER, by Thomas R. Harrison. 234 pages, diagrams, 6 × 9 in. New York, John Wiley & Sons, Inc., 1960. Price, $12.00.

The author's stated purpose in preparing the book was " . . . to present the subject of radiation pyrometry in enough theoretical and practical detail to enable operating engineers to arrive at definite answers to the many technical questions that arise in connection with . . . radiation pyrometers under the wide variety of conditions that are met with in industry." He has succeeded in accomplishing that purpose. The first three chapters deal with the principles of thermal radiation and their application to pyrometry. Chapter 4 is a detailed treatise on radiation pyrometers —types, characteristics, calibration, lens and mirror design. Chapter 5 is devoted to emittance corrections and Chapter 6 (62 pages) contains 21 tables and 4 charts of useful information such as millivolt equivalence tables, emissivity and emittance tables, emittance correction values, etc.

HYDROMAGNETIC CHANNEL FLOWS, by Lawson P. Harris. 90 pages, diagrams, 6 × 9 in. New York, John Wiley & Sons, Inc.; Cambridge, The Technology Press; 1960. Price, $2.75.

This short text gives the analyses for three flows of viscous, incompressible, electrically conducting fluids in high-aspect-ratio rectangular channels subjected to transverse magnetic fields. Subjects discussed thoroughly are turbulent flow in the presence of a d-c magnetic field and both laminar and turbulent flows in induction-driven flows. The methods used are analogous to the "semi-empirical" techniques of fluid mechanics in depending for their success upon the combined use of the basic hydromagnetic equations, a dimensional analysis, and the results of experiment. Principal results of this study include a scheme for correlating experimental friction-factor data and methods for computing mean-velocity profiles, mean magnetic-field distortion, current distribution, and turbulent-shear distributions in the d-c turbulent flow.

PUBLICATIONS RECEIVED

PHYSICS FOR STUDENTS OF SCIENCE AND ENGINEERING, PART II, by David Halliday and Robert Resnick. 510 pages, illustrations, 6 × 9¼ in. New York, John Wiley & Sons, Inc., 1960. Price, $6.00.

THE OTHER SIDE OF THE MOON, issued by the U.S.S.R. Academy of Sciences and translated by J. B. Sykes. 36 pages, illustrations, 8¼ × 10¾ in. New York, Pergamon Press, 1960. Price, $2.50.

MARINE RADIO FOR PLEASURE CRAFT, by Harold McKay. 160 pages, illustrations, 5½ × 8½ in. New York, Gernsback Library, Inc., 1960. Price, $2.95 (paper).

MECHANICS OF MATERIALS, by Archie Higdon, Edward H. Ohlsen and William B. Stiles. 502 pages, diagrams, 6 × 9 in. New York, John Wiley & Sons, Inc., 1960. Price, $7.75.

CALCULUS AND ANALYTIC GEOMETRY, by Walter Leighton. 483 pages, diagrams, 6 × 9 in. Boston, Allyn and Bacon, Inc., 1960. Price, $8.25.

SAFE HANDLING OF BODIES CONTAINING RADIOACTIVE ISOTOPES (A GUIDE FOR SURGEONS, PATHOLOGISTS, AND FUNERAL DIRECTORS). Handbook 65, National Bureau of Standards, 23 pages, 5½ × 7½ in. Washington, D. C., U. S. Department of Commerce, 1958. Price, $0.15 (paper).

TROISIÉME CONFERENCE INTERNATIONALE SUR LA PREPARATION DU CHARBON. Institut National de l'Industrie Charbonniere, 789 pages, illustrations, 8¼ × 11¾ in. Bruxelles, Editions Techniques et Scientifiques, n.d. English, German and French texts available (paper). Price not given.

THE NATURE OF SCIENCE, by David Greenwood. 95 pages, 4½ × 7¼ in. New York, Philosophical Library, 1959. Price, $3.75.

ARTIFICIAL EARTH SATELLITES, Volumes 1 and 2 (bound together), edited by L. V. Kurnosova and translated from the Russian. 107 pages, illustrations, 6 × 9 in. New York, Plenum Press, Inc., 1960. Price, $9.50.

THERMOCHEMISTRY FOR STEELMAKING, Volume I, by John F. Elliot and Molly Gleiser. 296 pages, tables, 8½ × 11 in. Reading (Mass.), Addison-Wesley Publishing Co., Inc., 1960. Price, $10.50.

ENGINEERING THERMODYNAMICS, by James B. Jones and George A. Hawkins. 724 pages, diagrams, 6 × 9 in. New York, John Wiley & Sons, Inc., 1960. Price, $8.50.

V.2
196

NATIONAL BUREAU OF STANDARDS NEWS

TRANSFERRING LIQUEFIED GASES BY PIPELINE

Liquefied gases, such as hydrogen, oxygen, nitrogen, and helium, are widely used as refrigerants in basic research in chemistry, physics, and engineering. More recently these liquids have also come to play an essential part in the propellant systems of rockets, missiles, and satellites. In general, the low-temperature liquids are transported from the liquefier to the experimental set-up in insulated, Dewar-type containers.

The current interest in liquefied gases, both in private industry and throughout Government, has led the National Bureau of Standards to take a second look at accepted methods for long-distance transportation of these fluids. The results of a recent theoretical study indicate that the pipeline method, now employed only in relatively short-distance transfers, can be extended to appreciable distances provided there is an awareness of the performance limitations introduced by undesirable cool-down characteristics. The design of such a system, which in many cases could undoubtedly increase efficiency, safety, and economy, is relatively straightforward. The Bureau's study, conducted by R. B. Jacobs of the Cryogenic Engineering Laboratory, located at Boulder, Colo., not only provides the theory for designing long-distance pipeline systems for low-temperature operation, but also gives information required for numerical computation.[1]

There are two categories of problems encountered in dealing with liquefied gas transfers: those that are predominantly of an unsteady nature and that occur prior to the transfer system's cooled-down condition; and those that can be considered as steady state and are applicable only after the system has been cooled to operating conditions. Only the latter are considered here.

The Bureau's study has a two-fold purpose: first to discover what equipment and techniques are required for the long-distance transfer of the liquids and whether these are available, and second to determine what losses will occur during these long-distance transfers and whether they are tolerable.

The primary component in the transfer's physical system is a length of insulated pipe—the transfer line. The state of the fluid is specified at both the inlet and outlet of this line. Either pumping systems or a process plant, such as a liquefier, can supply the energy required to force the fluid through the line. Pumping stations distributed along this transfer line may be more desirable than supplying all the necessary energy at the inlet. In this case, the system consists of a series of pipes with the fluid state specified at the inlet and outlet of each pipe. Since it is wasteful to carry unavailable energy (as sensible heat) through the system, intermediate stations along the line cool the liquid by one, or both, of two methods: refrigeration or flashing. If the

[1] NBS Circular 596 (December 1, 1958). Available from the Superintendent of Documents, U. S. Government Printing Office, Washington 25, D. C. Price, 30 cents.

See also "Cryogenic Piping Systems Design Considerations," by R. B. Jacobs, *Heating, Piping, and Air Conditioning*, Vol. 32, p. 129 (1960).

liquid is flashed to atmospheric pressure, its vapor is discarded or piped back to a liquefier and the remaining liquid, which is at its normal boiling point, is pumped through the next section of pipe.

The Bureau's conclusions are derived from a mathematical model based on the laws of conservation of energy, momentum, and mass. These laws were used along with verified assumptions and empirical data to yield a complete, unique model consisting of 12 equations which describe the flow characteristics of a transfer line. This flow model permits the design of a transfer system by relating all the pertinent design parameters: flow rate, line length, line diameter, pumping pressure, heat leak, and fluid properties.

Since the technical design of a system leaves economic and feasibility questions unanswered, the Bureau undertook further studies based upon the loss of liquefied gas in the system. There are four ways in which losses occur: loss at the pump due to energy dissipation in the pump container and cooling prior to line entry; flashing loss caused by heat leak and energy introduced by the preceding pump; cool-down loss which occurs when the system is cooled from ambient to operating temperature; and trapped-liquid loss—from liquid that cannot be removed at the end of the transfer. The Bureau formulated relations from which each loss can be computed and expressed them in a form which facilitates the required computations.

In using the mathematical formulations to design a particular system, empirical information must also be employed correctly. Equations for heat leak and fluid drag have therefore been included.

The Bureau is now in the process of experimentally verifying its theoretical conclusions on long-distance transfer systems. In addition, the operating data on some large systems currently being built will be used in an effort to put the present results on a firmer basis.

V.2
196

RECENT IMPROVEMENTS IN RADIO PROPAGATION PREDICTION SERVICE

The National Bureau of Standards has recently increased the scope and reliability of its radio propagation prediction service, widely used in long range radio communications. Practical experience with predictions over a period of years, together with a large volume of additional ionospheric and solar data obtained during the International Geophysical Year (IGY), has served to improve technique. Military and commercial users are now able to utilize NBS forecasts more effectively in communications to all parts of the world.

Much of the world's radio transmission depends on the reflection of electromagnetic waves by electrons, or ions, in that part of the atmosphere between 40 and 350 miles above the Earth's surface. The four levels of the ionosphere, D, E, F-1 and F-2, are not, however, sharply defined. They should not be thought of as distinct bands, but rather as amorphous groupings of particles, somewhat like sponges. These layers are differentiated by such characteristics as height, density and critical frequency (the highest frequency transmissions normally reflected by that level), and act very much like mirrors, reflecting high-frequency radio waves back toward the Earth. By alternate reflection from the ionosphere and the Earth's surface, radio signals in the range between 3 and 30 Mc can be transmitted over distances up to 4000 kilometers. The E- and F-regions are the most important to long distance communication.

Information about the constantly changing ionosphere is collected and analyzed by the Bureau's Central Radio Propagation Laboratory (CRPL) at Boulder, Colo. CRPL uses these data to predict optimum frequencies for communication between any two points in the world, based on vertical incidence soundings, for one-hop, F-2 transmissions for average layer height for a 4000 km distance over a great circle path. Predictions are published monthly as the CRPL-D series, and forecast for three months subsequent to publication.[2]

The characteristics of the ionosphere vary with geographic location. They also change with the season and with the sunspot cycle. Various factors responsible for the changes are not independent in their effects. There are also random fluctuations that are not completely explained. Thus, the variations of the ionosphere present a complicated geophysical phenomenon whose principles need to be well understood to achieve the best use of radio communications. To increase this understanding and to improve the predictions, the ionosphere was intensively studied during the International Geophysical Year.

A typical practical problem of radio communication may involve selecting a suitable frequency for use over a given path at a given time, within the limitations of the transmitting and receiving equipment. The frequency must be low enough to be reflected from the upper layers of the ionosphere to the receiving site; if too high a frequency is chosen, the signal will penetrate the ionosphere and be lost. Conversely, if too low a frequency is chosen, it will be largely absorbed in passing through the D-region of the ionosphere and will require excessive radiated power for satisfactory reception. Under certain conditions, this selection is easily accomplished; under other circumstances, no practical radio communication can be established. Consideration must also be given to the local conditions at the transmitting and at the receiving sites, as well as to limitations imposed by radio noise, fading, or interference. Effective use of the Bureau's propagation predictions aids greatly in establishing successful communications and in making efficient use of the radio spectrum.

International Geophysical Year

When the Bureau began the monthly predictions in 1942, they were based on information from only three vertical incidence ionospheric sounding stations. Since then, data available for analysis have increased vastly. Values were recorded by approximately 50 stations at the end of World War II, by 78 stations at the beginning of the IGY, and by 161 stations during IGY.

Before IGY, there were limited ionospheric data from South America, Africa, and Antarctica. To make predictions for these areas, in addition to the minimal direct data, data from the northern hemisphere were applied to conjugate points in the southern hemisphere for corresponding seasons of the year. Increased geographical coverage has resulted in more accurate predictions, as the need for interpolation and extrapolation of data has been greatly

[2] CRPL-D Series, "Basic Radio Propagation Predictions," available from the Superintendent of Documents, U. S. Government Printing Office, Washington 25, D. C. Annual subscription, $1.50. Use of the monthly predictions is explained in NBS Circular 462, "Ionospheric radio propagation," available from the above source. Price, $1.25.

reduced. At present, it is possible to use data from the stations currently in operation in the antarctic to add to our knowledge of the northern polar regions.

Not only were there a larger number of stations than ever before during IGY but more intensive data have been obtained at many stations which were active for longer periods of time. With such stations, the additional time coverage (2½ years, counting International Geophysical Cooperation 1959) has proved invaluable.

A further result of the increase in world-wide coverage has been the addition of two charts, covering an additional zone, to the monthly propagation predictions. The world map is now divided into four instead of three zones as before. The zones are centered approximately on the meridians 75° West, 20° East, 110° East, and 160° West, beginning with predictions for December 1958. This greater accuracy, taking into account longitudinal variations of critical frequency, was made possible largely by the increased number of stations reporting from Africa.

Sunspots

Sunspots strongly influence the ionosphere, and consequently affect terrestrial radio communications. Therefore, the International Geophysical Year was planned to take place during a period of intense solar activity to maximize the opportunity for studying sunspot effects. The 18 months beginning July 1957 turned out to be a fortunate choice. Not only did the current sunspot cycle reach its maximum in March 1958, but the maximum sunspot number of 201 for the 12-month running average exceeded by over 40 points the highest known previous maximum, occurring in the late 1770's. The average of the maxima of the previous recorded cycles was less than 100.

CRPL predictions are based on the assumption that there is a linear relationship between sunspot number and the critical frequency at which radio waves at vertical incidence just begin to penetrate the F-2 layer of the ionosphere. However, it was suspected, from limited evidence obtained during two previous sunspot maxima, 1937 and 1947, that the relationship would cease to be linear when the sunspot number rose high enough. That is, the ionization of the ionosphere would increase to saturation, and its maximum ion concentration would not change appreciably thereafter. Therefore, the number 150, which seemed to correlate with the saturated condition, was taken as the highest sunspot number in making predictions over the period 1956–1959, even though the actual number was higher at times. Subsequent comparisons of observed data with predictions showed that this choice had been correct. Statistical analysis of critical frequencies and sunspot number during this period gave further theoretical confirmation to the assumption.

During this period, maximum usable frequencies rose high enough to create many problems of interference, particularly on mobile communication circuits and on forward-scatter transmission in the 25–60 Mc range. One example was interference received by an oil company walkie-talkie in Texas from a police network in Rhode Island. This type of interference was most prevalent during daylight hours in the fall and winter.

Although the sunspot cycle had started to decline by the fall of 1958, the sunspot number was still well above all previous maxima. Interference was expected to be less, but it was worse at this time than for any other period

during the years 1956–1958. This excessive interference is attributed to the fact that the period from September 1958 through January 1959 was unusually free from ionospheric disturbances, which tend to lower critical frequencies. Magnetic disturbances seem to lag from one to three years behind the peak of solar activity. The increasing number of disturbed periods in the late summer of 1959 was therefore an indication that the solar activity cycle was definitely past its peak. However, during the first half of April 1960, the Earth's magnetic field was much disturbed.

Of great value to future prediction techniques would be another comprehensive study at vertical incidence during the next sunspot minimum, which will probably occur some time between 1965 and 1967. If such an effort can be made, only a minimum number of ionospheric sounding stations need be operated thereafter on a routine basis, and except for a few monitored ones, vertical sounding operations would only be needed for basic research purposes.

Experience with Predictions

In general, reports on the practical application of the predictions indicate that they have been somewhat conservative. Experience has shown that combinations of modes of propagation often occur giving longer path lengths, higher usable frequencies, and longer periods of satisfactory propagation. Since the predictions are limited in application to ionospheric propagation modes, the highest circuit efficiency is frequently a question of application rather than of the accuracy of the predictions themselves. For example, frequency usage on a North Atlantic circuit is often continued one or two hours past sunset, longer than would be expected from a study of the ionization over the great circle path. Because of continued ionization to the south of the path, reflection of the transmitted waves will continue many minutes after the ionization on the direct path has diminished to a level too low for communication purposes.

Furthermore, the type of transmitting antenna also influences the possible use made of the predictions. The reception from an omnidirectional antenna, such as is used by NBS radio station WWV, often results in the use of higher frequencies than might be expected. For instance, during periods when WWV might not be expected to be received in Ottawa, Canada, on a frequency greater than 5 Mc, it is often heard up to 20 Mc. The explanation is that energy from the WWV transmitter is reflected from the ionosphere in all directions; it is then reflected from the ground back to the ionosphere and after another reflection downward received at Ottawa. This form of back-scatter can effectively increase transmission distances by several thousand kilometers, making reception at the higher frequencies possible.

The many studies completed of various ionospheric phenomena and their variations with time, season, and solar activity afford an explanation of the successful extension of NBS predictions in actual communications. The facts derived from applications may explain why median maximum usable frequency may be a better estimate of optimum frequency on a given path than a percentage reduction figure derived from maximum usable frequency.

CURRENT TOPICS

Ceramic Telemetry Circuits.—A new approach to telemetry circuits, particularly in space applications, is suggested by General Electric in a line of small ceramic receiving tubes and "TIMM" ceramic circuits, which offer exceptional high gain and low noise performance at high frequencies. In addition, they are extremely tolerant of heat, variations in ambient temperatures, mechanical vibration and shock, and nuclear radiation.

"TIMM" circuits (thermionic integrated micro-modules) are said by General Electric to be ready for development into whatever specific circuitry is needed—and offer such possibilities as the equivalent of a 100-tube computer the size of a package of cigarettes. In TIMM circuits all necessary components such as resistors, capacitors, and tube elements are stacked, evacuated and sealed to form ceramic capsules which thrive on heat.

They are intended to be thermally insulated and operate principally on the heat generated within themselves —at near 600°C. They thus offer a practical solution to component density requirements equivalent to a million parts in each cubic foot.

Outstanding among ten types of ceramic receiving tubes is a developmental low-noise, broad-band triode amplifier which features transconductance of 40,000 micromohos.

Capable of operating at high temperatures, ceramic receiving tubes deliver the ultimate in high frequency performance in the small tube range without the use of blowers or other bulky, inefficient cooling systems which in effect tend to reduce the high theoretical component densities of solid state devices.

Ceramic tubes vary in diameter from about half that of a pencil to a half dollar, depending principally on power output capabilities. Low noise figures feature their UHF performance. The tubes are easy to mount, rugged, and provide flexibility in circuit design.

Defense and electronics industry personnel are becoming increasingly aware of the significance of the capabilities of these small ceramic tubes and their contributions to increased reliability of military equipment.

Of particular importance is the fact that ceramic receiving tubes are relatively immune to high level pulses of nuclear radiation which accompany nuclear weapons blasts. These high initial pulses do not harm human beings several hundred miles from their source, but can alter or disable some electronic devices either permanently or temporarily. In either case this could result in persons who operate and rely upon electronic equipment not even being aware of the fact that their transistorized equipment is not performing properly.

While transistorized digital computer circuits would fail when exposed to conditions equivalent in space to a one-megaton explosion as far as 1400 miles from the equipment, TIMM circuits would continue operating at least as close as 50 miles from the source of the initial gamma pulse caused by the explosion.

TIMM circuitry is the only electronic circuit concept known to be in-

herently resistant to the effects of pulse radiation, according to the General Electric Receiving Tube Department, Owensboro, Ky., where the circuits are being developed.

TIMM circuits at present are being made, in development, about the diameter of a cigarette. Lengths vary according to the needs and complexity of the circuit being built into the evacuated ceramic and metal capsule.

TIMM circuit development is employing two new concepts. In the first place, it was decided to provide a complete electronic circuit which has totally new components rather than trying to adapt present components to the modular circuit. Secondly, rather than trying to get rid of the heat always given off by electronic components as they do their work, TIMM circuits are designed to use this heat to operate.

By surrounding TIMM circuits with thermal insulating material, the heat created by the circuits while operating is retained. After an initial application of heat from an external source to start cathode emission, the circuits then heat themselves from the same electrical energy source by which they operate.

This source could be a solar converter in satellite applications.

Since TIMM circuits thrive on heat, all the power is used efficiently, no power is lost due to heat dissipation, and cooling devices needed for conventional electronic equipment are not required.

TIMM circuits can be built to fit almost any electronic requirements—and in the process eliminate some 20 per cent of the electronic components needed—and reduce by 40 per cent the electrical connections which often are the cause of an electronic equipment failure.

The inherent operating characteristics of TIMM circuits make them suited for use in future space probes similar to the sun-orbiting Pioneer V space vehicle. Since TIMM circuitry is resistant to radiation, no excess weight need be carried by the vehicle to shield the electronic equipment from space radiation.

Sunlight Communication System.—In some future year when a cruising space vehicle communicates with another space vehicle or its orbiting station, it may use a beam of light instead of conventional radio.

Not that radio will be inoperative under the airless conditions of space—rather the reverse—but there is reason to believe that communication by sunlight not only will be cheaper but will entail carrying much simpler and lighter equipment for certain specialized space applications.

The Air Research and Development Command's Wright Air Development Division is developing an experimental system that will collect sun rays, run them through a modulator, and direct the resultant light wave in a controlled beam to a receiver. There the wave will be put through a detector, transposed into an electrical impulse and be amplified to a speaker. Depending on the type of modulator used, either the digital (dot-dash) message or a voice message can be sent.

Since all components of the system are available, it is believed that a working model can be assembled by the end of this year. However, some components will require further development to make them adaptable to space flight.

The system works on the heliograph principle, but compares with that system in about the same ratio as an automobile does to a horse and wagon.

The sun's rays are collected by a Gassegraine type mirror system, that is, a concave mirror, which reflects the rays to a smaller mirror that directs them back, in a narrow beam, through a hole in the center of the original mirror. The collector will be directed toward the sun at all times by a sun tracking device.

The rays, passing through the small hole in the collector mirror, will be picked up by a series of movable mirrors, which will direct them to the modulator. This, in effect, is the shutter used in a heliograph. In the digital signal system, it will chop the light beam into dots and dashes, and in the voice system it will vary the intensity of the light into waves.

The modulated light beam then will be picked up by another set of movable mirrors, which will direct it to a transmitter mirror. There it will be beamed in a narrow band for transmission to the receiver.

The receiver will pick up a small fraction of the now widened beam, pass it through a series of mirrors to narrow and focus it, and then direct it into the detector. There the chopped up beam or the wave will be transposed into minute electrical impulses, which, when amplified, will be heard as sound, or seen as a written message.

One of the most difficult operations will be lining up the transmitter and receiver, but several methods of solving this are known.

Investigation of light communication was started about a year ago in search for a simpler, less cumbersome and jam-proof system than radio.

Among advantages of this method are the unlimited power source of the sun's rays, its relative simplicity and therefore higher reliability than radio; its possibly lighter weight, its virtual freedom from jamming, and avoidance of already overcrowded radio frequency bands.

The one major disadvantage of the system is visibility that would affect only communication with the Earth. Between space vehicles, between a space vehicle and a satellite station, and possibly between either of these and another planet, the system should be ideal. Communication with earth might be hampered by cloud covers, but it still could be utilized for transmitting voluminous scientific data on a non-priority basis.

Electrostatic Teleprinter.—The first production model of a 3000 word-per-minute teleprinter has been shipped by Burroughs Corporation to Kellogg Switchboard and Supply Co., Chicago, a Division of International Telephone and Telegraph Company. This machine called Burroughs High Speed Electrostatic Teleprinter, will be incorporated into the Strategic Air Command 465L control system by Kellogg. I.T.&T. is the prime contractor for designing and implementing the system.

Printing is accomplished with a row of solid state electronic print heads in matrix form that charge paper with electrostatic images. These images pick up powdered ink as the paper passes through an ink bath and are fixed into hard, permanent copy when the paper passes over heated pressure rollers. Eliminated are ordinary hammering type faces, typical of slow speed impact printers.

Since the paper never touches the print heads and the only moving parts are associated with the paper advance mechanism, the life and maintenance of the teleprinter revolutionizes conventional teleprinting equipment.

The teleprinter records alpha-numeric or symbolic characters at rates up to 300 characters per second on roll

paper 72 characters wide. It can be used as a direct readout from electronic computers, or operated from input signals from polka dot, punched or magnetic tapes, or telegraphic signals, and has capabilities of speeds up to 60,000 characters per minute.

Refrigerators for Space Ships.—General Electric scientists are working on a special type of refrigerator for those people who will be "dining out" in space. Under a contract from the Air Force's Wright Air Development Division, the Company's Missile and Space Vehicle Department at Philadelphia is studying how food reserves on space ships (appropriately shortened to FROST) can be preserved on space trips of long duration.

In designing food storage and preservation methods, scientists must consider the problems of the near-zero temperature of space, the possible effects of cosmic and other radiation, extreme heat during entry to the atmosphere of Earth or other planets and the near vacuum of space. They must also consider other factors inside the space vehicle such as available power supply, the type of water supply system used, and the effects of zero gravity. Along with these factors, the problems of weight and the amount of space required for the food preservation system are always primary considerations.

About nine pounds of supplies are needed for the daily nutritional requirements of one man, according to the scientists. The types and size of different food preservation methods required for space missions of long duration are now being investigated under the new Air Force contract. A working model of a proposed space "refrigerator" will be produced by General Electric as part of the food preservation study.

Engine for Satellite Space Power.—An engine no bigger than a milk bottle has been developed by Avco Corporation of Wilmington, Mass., to propel future vehicles in space. Called a 30-kilowatt arc jet engine, the miniature power plant already has operated continuously for two days and eventually will produce power to move satellites about in space for a period of several months.

The arc jet engine developed by Avco uses helium or hydrogen as a working liquid or fuel. It is regarded as a significant advance in the field of space propulsion, because, while it takes very large engines to place a satellite in orbit, very little power will move it about once it is on its assigned orbital path. Such engines as the 30-kilowatt arc jet would be used to make corrective changes in altitude or to rotate the satellite so that it could do its job more effectively. A satellite in orbit is in perfect balance and thus can be moved by relatively small amounts of continuous power over a period of hours or days. The small thrust of the engine imparting a continuous "push" on a satellite balanced in orbit around the earth would provide a steady acceleration and eventually very high speeds.

While operation of only a few minutes at a time would be sufficient to make corrective changes in altitude or position of the satellite, several hundred hours of operation might be needed to carry the satellite from low-altitude Earth orbit to around the moon in a constantly expanding spiral.

The engine itself weighs only 3½ lb., but as much as 1000 lb. of fuel and power supply might be needed for a long space operation.

Transparent Gallium Phosphide Crystals.—Scientists at Bell Telephone Laboratories are developing

methods of growing single crystals of a transparent semiconductor, gallium phosphide, and studying its properties. These studies will help in the basic understanding of high energy gap semiconductors, part of a field of research which has interested Bell Laboratories since well before its invention of the transistor in 1948. The material is transparent, making it possible to observe visually the differences which take place under varying conditions of doping and electron density.

Gallium phosphide belongs to the family of semiconductors referred to as "III-V compounds," which are composed of elements taken from each of the III and V groups of the periodic table. Knowledge gained in studying it should be applicable to other members of this group, some of which are already known to exhibit interesting properties. These include gallium arsenide and indium antimonide, which have shown promise in devices such as point contact and Esaki diodes. Many other III-V compounds have yet to be synthesized.

Although the semiconducting properties of gallium phosphide were first reported in 1954 by G. G. Folberth and F. Oswald in Germany, pure samples have been very difficult to prepare. One of the main problems is that it decomposes at temperatures near its melting point unless it is held under high pressures of gaseous phosphorous. The work at Bell has been aimed at producing single crystals of gallium phosphide by the floating zone melting technique. Special equipment had to be built to withstand the high pressure phosphorous atmosphere.

In the experimental equipment, an ingot of gallium phosphide is enclosed in a sealed quartz tube which is encased within a steel pressure chamber equipped with viewing ports. At the temperatures and pressures involved (1450° to 1500°C. and 10 to 50 atmospheres pressure), the walls of the quartz tube fog over, making observation of the molten zone difficult. This can be overcome to a large extent by heating the walls of the apparatus to above 950°C. Single crystals about $\frac{3}{4}$ in. long by $\frac{1}{4}$ in. square have been obtained at Bell Telephone Laboratories.

New Type 304 Stainless Steel.—High neutron absorption and excellent mechanical properties are combined in a new Type 304 stainless steel containing up to 2 per cent of the boron 10 isotope. The new alloy, available from The Carpenter Steel Co., Reading, Pa., is a big improvement over Type 304 with natural boron, and will be most useful for control rods, burnable poison and shielding for nuclear reactors. It offers improved design flexibility, while at the same time promises to increase the safety and reduce the maintenance costs of operating nuclear power plants.

For example, Type 304 with 0.2 per cent boron 10 has essentially the same nuclear properties as Type 304 with 2 per cent natural boron, but it has much better mechanical properties.

As the boron content in steel is increased, the tensile and yield strength increase while the ductility, impact strength and corrosion resistance decrease. Ordinarily, up to 2 per cent natural boron has to be added to give the steel a sufficiently high thermal neutron absorption cross section. This amount of boron cannot be added, however, without sacrificing important engineering properties.

Now, by adding a much smaller amount of enriched boron—0.2 per cent of the boron 10 isotope—it is possible to obtain mechanical prop-

erties and corrosion resistance that are almost as good as they are in the basic alloy. Although considerably less boron is added, there is no loss in nuclear properties.

If, on the other hand, an application requires greater neutron absorption than that obtainable with 2 per cent natural boron, up to 2 per cent of boron 10 may be added to the alloy. In such a case, the mechanical properties will not be as good as they are at a lower boron 10 level.

Selected Natural Diamond Abrasive.—Availability of a new type of natural diamond abrasive grit developed specially for use in resin-bonded grinding wheels has been announced by the Industrial Diamond Division of Engelhard Hanovia, Inc., of Newark, N. J. Designated "SND" for Selected Natural Diamond, the new material is claimed to provide up to 30 per cent greater efficiency than conventional diamond grit for certain operations in the grinding of cemented carbides.

Advantages of the new material stem from the geometric shape of the particles, which are predominantly needle-like or thin, flat plates. This is in contrast to the blocky particle shape of conventional natural diamond particles. Special boart crushing techniques have been developed and are being used to produce these special SND shapes.

Because of their shape, SND particles are held much more firmly in the resin bond than the more cubical shaped conventional diamond particles and are not pulled out of the wheel readily. Rather, the thinner particles break down gradually and evenly, continually providing sharp new cutting edges to the workpiece. In effect, SND thus provides many more cutting edges per carat than conventional material. An additional effect is reduced wheel loading and glazing.

Microscopic study of diamond particles recovered from the sludge and swarf after actual SND grinding evaluations shows these particles to be quite small and extremely uniform, confirming the fact that each original SND particle is almost totally consumed during the grinding.

Further experimental work resulted in the development of a special treatment for the SND grit which imparts still greater bond retention capability in resin wheels and increases particle friability. It has been found that both the treated SND grit and the untreated SND provide benefits under slightly different, but overlapping, carbide grinding conditions. As a result, the Industrial Diamond Division is making both of the new types available.

In general, the somewhat tougher untreated SND grit is recommended for feeds of between 0.0005 and 0.0015 in. and higher. The treated SND grit will provide the greatest efficiency increase at extremely light infeeds of under 0.0005 in., though benefits will still be significant in cuts of up to 0.001 in. For extremely heavy infeeds, 0.003 in. and over, the company recommends conventional natural diamond grit, which is also the preferred material for vitrified and metal bonded grinding wheels.

Radioactive Gas Used in Weather Studies.—A naturally-occurring radioactive gas has been used as a "tracer" to study how weather conditions affect air pollution, in experiments at the Argonne National Laboratory. The gas is radon—a product of radioactive disintegration of radium atoms in the soil. It has been estimated that every square mile of soil

to a depth of 6 in. contains approximately one gram of radium.

As radium steadily disintegrates by emitting alpha particles from its nucleus, it releases tiny amounts of radon through the soil into the atmosphere. The atmosphere contains on the average one part of radon to every billion billion (one sextillion) parts of air.

Scientists in the Meteorology Group of Argonne's Radiological Physics Division measured concentrations of radon in the air each hour during three observation periods totaling 72 hr.

This enabled them to observe how varying weather conditions affected the atmosphere's ability to cleanse itself of pollution. They also discovered new information concerning the atmospheric distribution patterns of radon. The test extended from ground level to 130 ft.

In the three-day project, scientists obtained air samples through a series of intake tubes mounted on top of a 130-ft. weather tower at the Argonne Meteorology Station. Air going into the tubes was measured, dehydrated and passed through charcoal filters, which trapped the radon. The filters were removed, processed in the laboratory and their radioactive content determined by means of a scintillation counter.

The next step was to compute how much radon was present per unit weight of air, by relating the amount of radon obtained to the volume of air from which it was extracted.

Scientists then related changes in the concentrations of radon in the air to changes in approximately thirty different meteorological variables. These variables are constantly measured at the Argonne Meteorology Station by special automatic recording instruments.

New Inorganic Glasses.—A new series of inorganic glass compositions, some of which are liquid at room temperature, were recently developed at Bell Telephone Laboratories. These glasses are composed of varying proportions of the elements arsenic, sulfur, and bromine.

Previous work at Bell Laboratories on compositions of arsenic, sulfur (or selenium), and thallium produced glasses which softened somewhat above room temperature, and became fluid enough to use for dip-coating at temperatures between 125 and 350°C. The substitution of iodine for thallium in the sulfur compositions later produced glasses with somewhat lower softening points, with several softening below room temperature. A few of these were highly fluid at temperatures below the boiling point of water.

The new ternary compositions with bromine show even lower softening and high-fluidity temperatures, with several as fluid as glycerin at room temperature. In spite of this, they are glassy solids at lower temperatures, show conchoidal fracture like typical silicate glasses, and exhibit no extensive ordering in their structure.

The new compositions appear to show the same characteristics as the previously reported low melting glasses. They are relatively stable toward acids, but are attacked by alkalies. They tend to hydrolize in water, with this tendency increasing in compounds showing lower softening points. The solids show good bulk resistivities, with values as high as 10^{16} ohm-cm. The resistivities of the liquid compositions are not as high, lying around 10^{8} ohm-cm.

Another interesting aspect of the new glasses lies in their optical transmission properties. They are all transparent, ranging in color from

V.2
196

ruby red to light amber. The indexes of refraction appear to range between 1.9 and 2.0, which is comparatively high for glasses.

Preparative techniques are relatively simple. The measured amounts of each of the constituents are weighed into a quartz tube with a pulled-down neck. The contents are then frozen in liquid nitrogen, the tube evacuated and sealed off. These sealed tubes or ampules are then encased in a simple pipe-bomb, and heated in an oven at about 400°C. for three to four hours. Simple agitation is sufficient to bring about reaction in the cases where glass formation occurs. The rate of cooling to room temperature does not appear to be critical.

These new glasses are products of early basic research efforts, and are not commercially available.

Computers Used to Orient Weather Satellite Pictures.—Specially designed computers produced by the Radio Corporation of America are helping scientists to orient the television pictures of cloud formations taken by the Tiros weather satellite. The computers—installed at Tiros ground stations at Kaena Point, Hawaii; Fort Monmouth, N. J., and RCA's Princeton Space Center—calculate the relative position of the satellite with respect to the sun at the time each TV picture is taken. Without such orientation, scientists would have difficulty interpreting what they saw.

The satellite itself contains a "north indicator" system which automatically measures the angle of the sun in relation to Tiros' spin axis. The system is made up of nine special solar sensing cells, arranged at points 40 degrees apart around the perimeter of the satellite. Each of the cells produces a distinctive pulse signal when it faces directly toward the sun as the satellite turns on its axis.

Each time a picture is taken, the satellite system transmits—along with the pictorial information—the distinctive pulses from the next two sensor cells to face the sun immediately after exposure. These pulses are separated from the picture signal and fed to the RCA sun-angle computer at the ground station. By measuring the time from picture exposure to the start of the first pulse, and the time between the two different pulses, the computer produces a figure for the exact angle of the sun in relation to the satellite's axis at the moment the exposure was made. This figure is then displayed with the picture to give scientists a guide line.

YOU CAN ADVANCE SCIENCE EDUCATION

Today, more than ever before in its 135-year history, there is vital need for The Franklin Institute to effectively promote education in science and technology. It is imperative that we meet this challenge by providing adequate educational opportunities in these fields. This requires vision, objective planning, and money. We have more than enough of the first two requisites, but far too little of the third.

Our programs are aimed at professional scientists and industry, as well as the lay public and young people seeking inspiration and guidance in choosing a career. The Institute's educational programs are impressive, for they begin with students in the early grades of our elementary schools and continue throughout an individual's professional or industrial life. With more funds at our disposal, the scope and vigor of these activities could be greatly increased.

The Franklin Institute is not richly endowed. It is a non-profit organization, depending for encouragement and support on an understanding public. Capable and conservative management assures wise administration of all funds.

Your gift or bequest, large or small, will be deeply appreciated and will be used effectively to broaden the Institute's educational usefulness. There is a warm satisfaction in giving financial support to an organization that has pioneered in, and is dedicated to, the advancement of science and technology through education.

When property is transferred, title should be in the name of The Franklin Institute of the State of Pennsylvania for the Promotion of the Mechanic Arts.

The Secretary of The Franklin Institute will gladly furnish you with additional information. Write to him at The Franklin Institute, Benjamin Franklin Parkway at Twentieth Street, Philadelphia 3, Pennsylvania.

V.
19

Journal
of
The Franklin Institute

MANAGING EDITOR, NANCY S. GLENN

Editor Emeritus, HENRY B. ALLEN

Associate Editors:

CHARLES B. BAZZONI
WILLIAM H. BOGHOSIAN
CHARLES L. CRITCHFIELD
RUPEN EKSERGIAN
VIRGIL MORING FAIRES
PAUL D. FOOTE
WALTER C. JOHNSON
ROBERT S. LIVINGSTON
C. E. K. MEES
F. J. MURRAY
W. F. G. SWANN
HUGH S. TAYLOR
MERLE A. TUVE

ARTHUR L. DAY, *Emeritus*

Committee on Publications:

HENRY B. ALLEN — LIONEL F. LEVY, *Chairman*, CLARENCE L. JORDAN — PHILIP H. WARD, JR.

Vol. 270 **AUGUST, 1960** **No. 2**

CONTENTS

New Gyros for Our Submarines ..... FREDERICK D. BRADDON 79

Harmonic and Subharmonic Response of an On-Off Control System to Sinusoidal Inputs ..... B. A. FLEISHMAN 99

Contribution to the Modified Z-Transform Theory ..... E. I. JURY 114

The First Electric Computer, A Magnetological Analysis .. GEORGE W. PATTERSON 130

The Franklin Institute
Bartol Research Foundation 138
The Franklin Institute Laboratories 143

Book Reviews 146

Publications Received 148

Book Notes 148

National Bureau of Standards News 152

Current Topics 156

Published monthly by The Franklin Institute of the State of Pennsylvania, Benjamin Franklin Parkway at Twentieth St., Philadelphia 3, Pennsylvania. Printed by Lancaster Press, Inc., Prince and Lemon Sts., Lancaster, Pennsylvania.

Subscription rates: Domestic, $10.00 a year; Foreign, $11.00 a year. Single current issues, $1.00. Publication of reprints for sale was discontinued in July, 1959. All inquiries regarding subscriptions, change of address, and sale of back issues, should be addressed to the Philadelphia office.

Indexes to the semi-annual volumes of the JOURNAL are published in the June and December numbers.

Copyright © 1960, by The Franklin Institute of the State of Pennsylvania.

THE FRANKLIN INSTITUTE
LABORATORIES FOR RESEARCH AND DEVELOPMENT

announce opportunities for:

ELECTRICAL ENGINEERS

SYSTEMS ENGINEERS

ORGANIC CHEMISTS

OPERATIONS RESEARCH SPECIALISTS

MECHANICAL ENGINEERS

EXPERIMENTAL PSYCHOLOGISTS

METALLURGISTS

INDUSTRIAL ENGINEERS

PHYSICISTS

APPLIED MATHEMATICIANS

to work on challenging problems in research

Send complete resume to:

Mr. John E. Christ, *Director of Personnel*
THE FRANKLIN INSTITUTE
Philadelphia 3, Pa.